



<http://www.diva-portal.org>

Postprint

This is the accepted version of a paper presented at *International Conference on Global Software Engineering*.

Citation for the original published paper:

Britto, R., Freitas, V., Mendes, E., Usman, M. (2014)

Effort Estimation in Global Software Development: A systematic Literature Review.

In: *Proceedings of the 2014 9th IEEE International Conference on Global Software Engineering* (pp. 18-21).

<http://dx.doi.org/10.1109/ICGSE.2014.11>

N.B. When citing this work, cite the original published paper.

Permanent link to this version:

<http://urn.kb.se/resolve?urn=urn:nbn:se:bth-11144>

Effort Estimation in Global Software Development: A Systematic Literature Review

Ricardo Britto
Software Engineering
Department
Blekinge Institute
of Technology
Sweden
ricardo.britto@bth.se

Vitor Freitas
Federal University
of Juiz de Fora
Brazil
vfreitas@ice.ufjf.br

Emilia Mendes
Software Engineering
Department
Blekinge Institute
of Technology
Sweden
emilia.mendes@bth.se

Muhammad Usman
Software Engineering
Department
Blekinge Institute
of Technology
Sweden
muhammad.usman@bth.se

Abstract—Nowadays, software systems are a key factor in the success of many organizations as in most cases they play a central role helping them attain a competitive advantage. However, despite their importance, software systems may be quite costly to develop, so substantially decreasing companies' profits. In order to tackle this challenge, many organizations look for ways to decrease costs and increase profits by applying new software development approaches, like Global Software Development (GSD). Some aspects of the software project like communication, cooperation and coordination are more challenging in globally distributed than in co-located projects, since language, cultural and time zone differences are factors which can increase the required effort to globally perform a software project. Communication, coordination and cooperation aspects affect directly the effort estimation of a project, which is one of the critical tasks related to the management of a software development project. There are many studies related to effort estimation methods/techniques for co-located projects. However, there are evidences that the co-located approaches do not fit to GSD. So, this paper presents the results of a systematic literature review of effort estimation in the context of GSD, which aimed at help both researchers and practitioners to have a holistic view about the current state of the art regarding effort estimation in the context of GSD. The results suggest that there is room to improve the current state of the art on effort estimation in GSD.

I. INTRODUCTION

Many software development companies have tried to increase profits by improving the time-to-market of their products, reducing costs by hiring people from countries with cheaper work-hours and defying the "clock" by running the projects during 24 hours. As a result, we can see a great number of software development projects performed globally, distributed in many different sites and normally located in different countries. This distribute setting of managing a software project is called Global Software Development (GSD).

The advantages related to GSD can only be achieved if projects are well managed. There are many challenges related to managing GSD projects. We could remark the following key challenges [1], [2]:

- **Communication** Different languages used by each site of a GSD project hinder the execution of the work.
- **Trust** Cultural variation complicates the process of trusting between each site.

- **Coordination over distance** - Different time zones makes harder the process of hand off between sites.

Those challenges affect directly the effort estimation of a project, which is one of the critical tasks related to the management of a software development project [3].

There are many studies related to effort estimation methods/techniques for co-located projects [4]–[6]. Nevertheless, existing evidence suggests that effort estimation methods/techniques and cost drivers used in the context of co-located projects may not be readily applicable to global software development projects [7]–[10].

There are many systematic literature reviews (SLR) regarding GSD. Kroll *et al.* [11] identified areas within Software Engineering have been investigated regarding GSD. Schneider *et al.* [12] performed a SLR to categorize solutions in the GSD context by using process areas. Jalali *et al.* [13] and Hossain *et al.* [14] performed systematic literature reviews looking at the use of agile methods in GSD. Da Silva *et al.* performed two SLRs on project management in the context of GSD [15], [16]. Prikładnicki *et al.* [17] identified papers that describe process models for GSD in the context of overseas outsourcing. Mishra *et al.* developed a SLR to find the current research and practice of quality management in GSD projects [18]. Smitte *et al.* [19] investigated empirical evidence in GSD-related research literature. Nurdiani *et al.* [20] discuss risks and strategies to mitigate those risks in the context of GSD projects. Monasor *et al.* [21] present a SLR regarding methodologies for training and teaching students and developers to the needs of GSD projects. Nidhra *et al.* [22] present the state of the art and the state of the practice regarding knowledge transfer in GSD. Matusse *et al.* [23] investigated evidences about the existence of metrics and indicators that are specific to GSD. Many of these previous SLRs focused on the challenges and solutions regarding communication, coordination and cooperation problems in GSD projects [24]–[27].

Nevertheless, to the best of our knowledge, there is no SLR regarding effort estimation in the context of GSD. Since nowadays GSD plays an important role to the software industry, a SLR on effort estimation in the context of GSD is deemed quite important as its results on the state of the art in that

field can inform both researchers and practitioners. Therefore, the goal and main contribution of this paper is to present a SLR of effort estimation in the context of GSD. In doing so, research gaps that can be the focus of future work were also identified.

The remainder of this paper is organized as follows: Section 2 details the SLR, followed by the presentation and discussion of its results in Sections 3, 4 and 5 respectively. Finally, Section 6 presents the conclusions and view on future work.

II. SYSTEMATIC LITERATURE REVIEW

This section details the SLR, which includes the research questions that guided this SLR as well as the procedures employed to perform the entire SLR process. The SLR detailed herein followed the guidelines defined by Kitcheham and Charters [28]. The protocol was developed jointly by the first, third and fourth authors.

A. Research questions

The research questions were formulated by using the PICOC criteria [29]. Note that we did not include the comparison attribute because our focus is not to compare interventions. The PICOC is as follows:

- **Population** - Global software development projects.
- **Intervention** - Effort estimation methods/techniques/size metrics/cost drivers.
- **Outcomes** - The accuracy of the effort estimation methods/techniques.
- **Context** - Any possible study, as long as it is an empirical study within the context of GSD will be considered.

Therefore, the addressed research questions are:

- **Question 1** - What methods/techniques have been used to estimate effort in GSD?
 - **1a** - What metrics have been used to measure the accuracy of effort estimation methods/techniques in GSD projects?
 - **1b** - What are the accuracy levels for the observed estimation methods?
- **Question 2** - What effort predictors (cost drivers/size metrics) have been used to estimate effort in GSD?
- **Question 3** - What are the characteristics of the datasets used for effort estimation in GSD?
 - **3a** - What are the domains represented in the dataset (academia/industry projects)?
 - **3b** - What are the types represented in the dataset (single-company/cross-company)?
 - **3c** - What are the application types represented in the dataset (web-based/traditional)?
- **Question 4** - What are the used sourcing strategies (offshore outsourcing/offshore insourcing)?
 - **4a** - What are the countries involved?
 - **4b** - How many sites are involved?
 - **4c** - How many sites per country are involved?
 - **4d** - Which topologies (centralized/distributed) have been used regarding the effort estimation process?

- **4e** - How did each of the sites participate in the effort estimation process (knowledge/data/both)?

- **Question 5** - Which activities were considered in the effort estimation process?

B. Search strategy

After defining the research questions, we set up a strategy to define a search string and, consequently, identifying the primary studies. To avoid researcher bias, we used the following procedure to define the search string used in this paper:

- 1) Analyze the questions and identifying the main words in terms of population, intervention and outcome;
- 2) Analyze collected relevant papers and checking the keywords;
- 3) Point out alternative spellings and synonyms for major terms, if existing;
- 4) Connecting alternative spellings and synonyms by using the Boolean OR;
- 5) Linking the main terms from population, intervention and outcome by using the Boolean AND;
- 6) Evaluate the resultant search string and perform adjustments by using the quasi-gold standard [30].

The steps one to five were performed by the first author and the last one was performed jointly by the first and fourth authors.

As a result, we obtained the following search string:

(effort OR cost OR resource OR size OR metric OR measure OR measurement) AND (estimation OR estimating OR estimate OR prediction OR predicting OR predict OR assessment OR forecasting OR forecast OR calculation OR calculate OR calculating OR calculation OR sizing OR measure OR measuring) AND ("global software development" OR "global software engineering" OR "globally distributed development" OR "globally distributed work" OR "distributed software development" OR "distributed software engineering" OR "distributed development" OR "multi-site development" OR "multisite development" OR "multi site development" OR "geographically distributed software" OR "collaborative software engineering" OR "collaborative software development" OR "dispersed software development" OR "global software teams" OR "distributed teams" OR "spread teams" OR "dispersed teams" OR "global teams" OR "virtual teams" OR "offshore outsourcing" OR "offshore software development")

C. Search process

Once the search string was defined, the search process took place and was divided in two phases: initial search phase and secondary search phase.

1) *The initial search phase:* The initial phase involved identifying primary sources and searching primary studies through those sources by using the search string defined in Subsection II-B.

The used primary sources are listed in Table I, as well as the number of search results and number of relevant papers. We selected those primary sources based on venues where

the literature related to GSD and effort estimation is mostly published.

The search process was performed using titles and abstracts. In addition, results were limited to peer-reviewed conference papers and journal articles, published between 2001 and 2013. The starting year was set to 2001 because that was the year when the term Global Software Development was coined [31].

TABLE I
SUMMARY OF SEARCH RESULTS

Database name / Search engine	Search results	Selected articles
Scopus	299	17
IEEEExplore	69	9
ACM Digital Library	59	7
ScienceDirect	26	0
Compendex	85	13
Inspec	108	10
Web of Science	63	9
Total	709	65
Without duplicates	379	24

2) *The secondary search phase:* The secondary search phase was carried out in order to maximize the chances that all important studies were included in this SLR. As part of this phase, all the references cited in the papers gathered in the first phase were checked. This phase was performed jointly with the study selection procedure.

D. Study selection

The study selection was carried out by applying the following inclusion and exclusion criteria:

- Inclusion criteria:
 - 1) Studies that present effort estimation models/methods/metrics for GSD **AND**;
 - 2) Studies that have empirical evidence **AND**;
 - 3) Studies described in English **AND**;
 - 4) Studies reported in peer reviewed workshop **OR** conference **OR** journal **OR** are reported in a technical report **OR** thesis found as a result of the secondary search phase.
- Exclusion criteria:
 - 1) Studies that do not present effort estimation models/methods/metrics for GSD **OR**;
 - 2) Studies without empirical evidence **OR**;
 - 3) Studies that are not written in English;

The screening process, which was used to select primary studies, was conducted in two stages. The first stage, performed by the first three authors, comprised reading studies' titles and abstracts. A total of 24 studies were judged as relevant (Table I). All the works selected in that stage were identified by a letter (S) and a number ranging from 1 to 24.

We downloaded the full-text of the studies selected in the first stage, in order to perform the second stage of the studies' selection process. The first author was responsible to download the full-texts of the studies retrieved from Scopus, IEEEExplore,

Web of Science and Inspec. The second author downloaded the full-texts of the studies retrieved from Compendex, and ACM digital library. No full-text was downloaded from ScienceDirect, since no study from that database was considered relevant.

In the second stage, we read the full-text of the 24 papers selected in the previous stage. The secondary search phase (II-C) was performed concurrently. The final list of the first three authors were compared and discussed until consensus was reached.

As a result of the secondary search phase, the first author found an additional study. After a discussion with the other authors, that study was included, thus making it a total of 25 studies to be screened. This 25th study was identified as D1, to be differentiated from the other 24 studies retrieved in the initial search phase.

As a result of the second stage of the study selection, the authors consensually decided to exclude 17 studies, due to the following reasons:

- Nine studies were excluded because they did not focus on effort estimation models, methods, size metrics or cost drivers in the GSD context (Exclusion criterion 1).
- Eight studies were excluded because they did not provide empirical evidences to support their findings (Exclusion criterion 2).

Thus, by the end of the study selection phase, 8 papers passed our inclusion criteria and were assessed in terms of quality (Subsection II-E).

E. Study quality assessment

In order to evaluate the quality of the selected studies, we developed a questionnaire with 13 questions. The quality score for a paper could range from 0 to 13. The higher the score, the better the quality of the evidence provided by the paper. The questionnaire was developed by using the guidelines defined by Kitcheham and Charters [28].

Each question could be answered using YES (1.0), NO (0.0) or PARTIALLY (0.5). If a question was answered with PARTIALLY, it means that the work addresses the question in a correct way, but the answer could not be confirmed by using the content of the paper. Any study which scored 3.25 or below (first quartile) was excluded.

The questionnaire had the following questions:

- Are the research aims clearly specified?
- Was the study designed to achieve these aims?
- Are the prediction techniques used clearly described and their selection justified?
- Are the variables considered by the study suitably measured?
- Are the data collection methods adequately detailed?
- Is the data collected adequately described?
- Is the purpose of the data analysis clear?
- Are the statistical techniques used to analyze the data adequately described and their use justified?
- Are the results discussed, rather than only presented?

- Do the researchers discuss any problems with the validity/reliability of their results?
- Are all research questions answered adequately?
- How clear are the links between data, interpretation and conclusions?
- Are the findings based on multiple projects?

The first and the second authors of this work have performed the quality assessment for the 8 selected primary studies. The third author evaluated one of the 8 selected studies. The results were discussed until consensus was reached. The final scores can be seen in Table II-E. Three studies were excluded due to their low quality score.

TABLE II
FINAL SCORES FOR THE EVALUATED WORKS

Study ID	Score
S7	10.5
S10	7
S11	10.5
S13	3
S15	13
S20	5
S24	3
D1	3

The lists with the included and excluded studies after the study selection and quality assessment are shown in Table II-E.

TABLE III
INCLUDED AND EXCLUDED WORKS AFTER THE STUDY SELECTION AND QUALITY ASSESSMENT

Included works	Excluded works
S7 [9], S10 [32], S11 [38], S15 [7], S20 [8]	S1 [33], S2 [34], S3 [35], S4 [36], S5 [37], S6 [39], S8 [10], S9 [40], S12 [41], S13 [42], S14 [43], S16 [44], S17 [45], S18 [46], S19 [47], S21 [48], S22 [49], S23 [50], S24 [51], D1 [52]

F. Data extraction

The first and the second authors of this paper have performed the data extraction for all the five final selected studies. The third author performed the extraction procedure for one of those five studies, thus to compare the quality of the extractions.

All the extracted values were compared and discussed until consensus was reached. The performed comparison was very important since some values were noticed just by one of the authors.

III. RESULTS

After the data extraction, the extracted data were consolidated in 12 tables so to make it easier to understand the relationship between data and the research questions.

The percentage column, which is part of all tables of this section (except for Table VI), indicates the percentage of studies out of the total number of studies (five).

A. Question 1

Data extraction for Question 1 looked at what methods which have been used for estimating effort in GSD, the accuracy metrics to evaluate the methods and also the calculated accuracy values. Those data were synthesized in Tables IV, V and VI.

Table IV shows the identified effort estimation methods/techniques, representing three different approaches: i) expert based approaches (Delphi [53], planning poker [54], expert judgment [55] and expert judgment based on ISBSG database [56]); ii) algorithmic based approaches (COCOMO II [55] and SLIM [57]); and iii) an artificial intelligence approach (Case-based reasoning [55]).

One of the primary studies did not investigate directly any effort estimation approach, but rather focused on the cost drivers. Another primary study used function points and use case points as surrogates for effort estimation techniques. However, both are size metrics. The primary study S20 presented an approach based on linear regression but that approach was not empirically evaluated. So, from S20 we only extracted cost drivers as they were empirically validated.

TABLE IV
IDENTIFIED EFFORT ESTIMATION METHODS

Estimation method	Approach type	Study ID	Percentage (%)
COCOMO II	Algorithmic	S11	20
SLIM	Algorithmic	S11	20
Case-based reasoning	Artificial Intelligence	S15	20
Expert judgment	Expert	S7	20
ISBSG based expert judgment	Expert	S11	20
Delphi	Expert	S7	20
Planning poker	Expert	S7	20
Function point count	-	S7	20
Use case point count	-	S7	20
No estimation approach	-	S10	20

Table V shows the accuracy metrics used in the primary studies. As we can see, just one of the primary studies presented the accuracy of the proposed effort estimation approach based on the most used accuracy metrics MMRE, MdMRE and Pred(25) [55].

The MRE (Magnitude of Relative Error) is calculated by considering the absolute difference between the actual and estimated effort, relative to the actual effort. The MMRE and MdMRE are respectively the mean and median MRE. Pred(25) is the the percentage of estimates with a MRE of 25% or less.

The primary study S11 used just the difference between the actual and the estimated efforts (deviation). The primary study S10 did not use any accuracy metric, since that work just focused on the cost drivers. The studies S7 and S20 did not evaluate the considered effort estimation approaches in terms of accuracy metrics.

Table VI shows the measured accuracy values for each observed effort estimation technique/method. As above-mentioned, only study S15 used MMRE, MdMRE and

TABLE V
IDENTIFIED ACCURACY METRICS FOR THE OBSERVED EFFORT ESTIMATION METHODS

Accuracy metric	Study ID	Percentage (%)
Deviation	S11	20
MMRE	S15	20
MdMRE	S15	20
Pred(25)	S15	20
No accuracy metrics	S7, S10, S20	60

Pred(25) to evaluate the accuracy of the presented effort estimation approach, thus making it difficult to assess how good were the estimation techniques put forward in studies S7 and S11.

In S11, three projects were used to evaluate three different approaches. The deviation between the actual and the estimated effort was calculated for each approach in the context of each project.

The ISBSG database [56], used in S11, provides three different estimation values: the lower time to perform a project; the time in which a project is more likely to be performed; and the maximum time to finish a project. So, in S11, for each project three deviations were calculated using the actual effort and the three above-mentioned values: Lower (L), Expected (E) and Upper (M).

When applying the thresholds suggested by Conte *et al.* [58] (good MMRE if calculated value $\leq 25\%$, good MdMRE if calculated value $\leq 25\%$ and good Pred(25) if calculated value $\geq 75\%$) to assess the effort prediction accuracy detailed in S15, results suggest that the prediction technique proposed in S15 provided good estimates.

It is impossible to compare the estimation accuracy between the effort estimation approaches identified in our SLR, since most of the primary studies did not use any prediction accuracy metric. In addition, the two studies that employed some accuracy measure did not use the same one.

B. Question 2

Data collection for Question 2 looked for data regarding the used cost drivers and size metrics. The extracted data is presented in Tables VII and VIII, where all the predictors that do not represent size measures were grouped as cost drivers. Note that whenever the same cost factor presented different names in primary studies, we chose the name we believed to be the most suitable and added it to Tables VII and VIII.

Table VII shows that there is a wide range of cost drivers in the selected primary studies, where only a single study (S7) did not show any cost driver.

Most of the cost drivers that seem to impact GSD projects are related to the specific factors of globally performed software projects, like differences in culture, language and time zone. The cost drivers that were used the most were "Time zone" (80% of the primary studies), "Language and cultural differences" (60% of the primary studies), "Communication" (40% of the primary studies) and "Process model" (40% of the primary studies).

TABLE VI
IDENTIFIED ACCURACY LEVELS FOR THE OBSERVED EFFORT ESTIMATION METHODS

Estimation method	Study ID	Accuracy
COCOMO II	S11	Project 1: 2.26; Project 2: 1.35; Project 3: 1.53
SLIM	S11	Project 1: 0.36; Project 2: 1.55; Project 3: 2.03
ISBSG	S11	Project 1: 0.68 (L), 3.44 (E), 12.65 (U); Project 2: 1.96 (L), 2.42 (E), 12.24 (U); Project 3: 0.3 (L), 4.23 (E), 14.53 (U).
Case-based reasoning	S15	MMRE:15.99%; MdMRE: 11.67%; Pred(25): 84.12%
Expert judgment	S7	no accuracy calculated
Planning poker	S7	no accuracy calculated
Delphi	S7	no accuracy calculated
Function point count	S7	no accuracy calculated
Use case point count	S7	no accuracy calculated

TABLE VII
IDENTIFIED COST DRIVERS

Cost driver	Study ID	Percentage (%)
Time Zone	S10, S11, S15, S20	80
Language And Cultural Differences	S10, S20, S15	60
Communication	S11, S20	40
Process Model	S15, S20	40
Communication Infrastructure	S10	20
Communication Process	S10	20
Travel	S20	20
Competence level	S10	20
Requirements Legibility	S10	20
Process Compliance	S10	20
Response Delay	S11	20
Unrealistic Milestones	S11	20
People Interesting	S11	20
Trust	S11	20
Clients Unawareness	S11	20
Shared Resources	S11	20
Team Structure	S11	20
Work Pressure	S11	20
Work Dispersion	S15	20
Range of Parallel-sequential Work Handover	S15	20
Client-specific Knowledge	S15	20
Client Involvement	S15	20
Design and Technology Newness	S15	20
Team Size	S15	20
Project Effort	S15	20
Development Productivity	S15	20
Defect Density	S15	20
Rework	S15	20
Reuse	S15	20
Project Management Effort	S15	20

Table VIII lists the size metrics employed in the primary studies, showing that function points and lines of code are the most used size metrics in the context of effort estimation for GSD. One of the studies also employed use case points and user story points as possible size metrics. The primary study S10 did not show any size metrics because it focused solely on cost drivers.

The use of lines of code as size metric seemed somewhat surprising to us given that this metric is really very difficult to forecast early in the development life cycle [59]. Such difficulties were one of the main motivations as to why function points were proposed so many years ago.

It is interesting to notice that just length metrics (lines of code) and functionality metrics (function points, use case points and story points [60]) were identified in the selected primary studies, despite other possible types of size metrics (e.g. complexity metrics, like cyclomatic complexity [61]).

TABLE VIII
IDENTIFIED SIZE METRICS

Size metric	Study ID	Percentage (%)
Function points	S7, S11, S20	60
Lines of code	S11, S15, S20	60
Use case points	S7	20
Story points	S7	20
No size metric used	S10	20

C. Question 3

Data collection for Question 3 looked at domains and types regarding the data used to evaluate the proposals in each primary study. That question also looked at the evaluated application types. The data were arranged in Tables IX, X and XI.

Table IX presents the domain relating to the type of dataset used in the primary studies' investigation. It shows that all the selected primary studies employed industrial data to evaluate and/or making conclusions about the effort estimation approaches and/or predictors (size metrics and cost drivers).

TABLE IX
IDENTIFIED DATASET DOMAINS

Domain	Study ID	Percentage (%)
Industry	S7, S10, S11, S15, S20	100
Academia	none	0

Table X presents the identified dataset types in the primary studies. Two of the primary studies (S10 and S15) used single-company data (data that comes from a single company). Study S11 used cross-company data, since the ISBSG database was used to estimate effort in that study. Studies S7 and S20 did not state the type of the dataset used.

Table XI presents the identified application types used in the primary studies. We considered only two types of applications during the data extraction - traditional and Web-based applications. In this work, a traditional application means an

TABLE X
IDENTIFIED DATASET TYPES

Type	Study ID	Percentage (%)
Single-company	S10, S15	40
Cross-company	S11	20
Not stated	S7, S20	40

application which does not require Web infrastructure. Only one study (S11) stated the types of the application used in the study, which included both traditional and web-based applications.

TABLE XI
IDENTIFIED APPLICATION TYPES

Types	Study ID	Percentage (%)
Web-based	S11	20
Traditional	S11	20
Not stated	S7, S10, S15, S20	80

D. Question 4

Data collection for Question 4 looked at specific aspects of Global Software Development. The extracted data is presented in Tables XII, XIII, XIV and XV.

Table XII presents the identified sourcing strategies in the primary studies. During the data extraction, we considered three options to address the question about sourcing strategies:

- **Offshore outsourcing** - A company moves the software development to an external third party abroad [62].
- **Offshore insourcing** - A company moves the software development to a branch established abroad [62].
- A combination of both above-mentioned strategies.

Table XII shows that just the offshore insourcing strategy was considered by the selected primary studies (S7, S10, S11, S15). One of the works did not inform which kind of sourcing strategy was used (S20).

TABLE XII
IDENTIFIED SOURCING STRATEGIES

Sourcing strategy	Study ID	Percentage (%)
Offshore insourcing	S7, S10, S11, S15	80
Offshore outsourcing	none	0
Both	none	0
Not stated	S20	20

Table XIII presents the identified number of involved countries in the collected works. In the selected primary studies, at least three countries were identified as involved in the context of a GSD project.

Table XIV presents the name of the identified countries, thus showing a wide range of countries. The USA, China, India and UK are the four countries mostly involved in GSD projects within the context of this SLR.

All the selected studies that presented this information had at least a site in a developed country and a site in developing

TABLE XIII
IDENTIFIED NUMBER OF INVOLVED COUNTRIES

Number	Study ID	Percentage (%)
3	S11	20
7	S10	20
10	S7	20
Not stated	S15, S20	40

country. Since one of the main reasons for using GSD is the cost-reduction, it is not surprise that those countries were identified. USA and UK have higher costs, specially regarding human resources, compared to countries like India [63].

TABLE XIV
IDENTIFIED COUNTRIES

Name	Study ID	Percentage (%)
USA	S7, S10, S11	60
UK	S7, S11	40
India	S7, S10	40
China	S7, S10	40
Malaysia	S7	20
Japan	S7	20
Taiwan	S7	20
Ireland	S7	20
Brazil	S7	20
Slovak Republic	S7	20
Finland	S10	20
Germany	S10	20
Norway	S10	20
Sweden	S10	20
Pakistan	S11	20
Not stated	S15, S20	40

Table XV presents the identified number of sites in the primary studies. A primary study should have at least two participating sites, each placed in a different country, to be characterized as within the context of GSD.

We observed a maximum number of four countries in a GSD project. Study S7 did not state the exact number of sites, and studies S15 and S20 did not report the number of sites involved.

TABLE XV
OVERALL NUMBER OF SITES

Number of sites	Study ID	Percentage (%)
2	S10, S11	40
more them 2	S7	20
3	S10	20
4	S10	20
Not stated	S15, S20	40

Question 4d deals with the effort estimation process topologies. In this work, the topology of the effort estimation process is the way the work regarding the estimation process is performed by the involved sites. There are three possible topologies:

- **Centralized** - All the inherent effort estimation calculations are centralized in just one site and the remaining sites just supply knowledge and/or data to the central site.

- **Distributed** - Each site of a GSD project performs their own effort estimation process independently. The estimates only relate to what the site is doing as part of the project.
- **Hybrid** - All the early effort estimation is performed as in a centralized topology, but late effort estimates can be obtained in a distributed fashion. We consider early effort estimation the one performed just after the requirement analysis. The late effort estimation process is the one performed after an implementation iteration.

Question 4e deals with the roles for sites in the effort estimation process. As above-mentioned, a site can be involved in the effort estimation process as a knowledge/data supplier and/or as the responsible for performing the inherent calculations.

We identified that none of the selected primary studies documented data about the aspects topology and the roles for a site in the effort estimation process. For this reason, we did not address Questions 4d and 4e.

E. Question 5

Data extraction for Question 5 looked at the activities of a GSD project considered in the effort estimation process. Herein we considered generic activities embraced by most of the well-known life cycles, e.g. Requirements and Testing.

The main goal in extracting this data was to identify the life cycle activities considered the most relevant as far as effort estimation is concerned, within the context of GSD projects. However, it was not possible to address this question, since none of the selected primary studies documented this aspect either.

IV. DISCUSSION

In order to facilitate the reading of our observations regarding the extracted data, we divided our comments according to the SLR research questions.

Regarding Question 1 we have identified that:

- The used effort estimation approaches are well-known in the context of co-located projects. Those approaches have been used in the context of globally distributed projects without any kind of adaptation in the mechanisms of each approach.
- There is no standard effort estimation approach, since all the primary studies presented different solutions to the observed problem. However, it seems that expert-based approaches are the ones used the most by the practitioners.
- Software engineering works with imprecise and uncertain knowledge [64]. However, none of the primary studies treated explicitly the uncertainty regarding the data usually used to perform the effort estimation process.
- Since just one of the selected studies employed commonly used accuracy metrics, it was impossible to compare the primary studies in terms of prediction techniques' estimation accuracy.

Regarding Question 2 we have observed that:

- The primary studies showed a wide range of cost drivers where most of them were related with the GSD specific issues, like cultural, language and time zone differences.
- None of the studies discussed deeply about the listed cost drivers.
- The observed primary studies show a trend in applying original co-located effort estimation approaches in GSD projects with cost drivers which reflect the specific problems of globally distributed projects.
- Solely length and functionality size metrics were identified in the studies, where the metrics used the most are lines of code and function points.

Regarding Question 3 we have identified that:

- The extracted data showed that only industrial data has been considered to evaluate the proposed effort estimation approaches.
- Single-company data appears to be the most used type.
- Most of the primary studies did not explicitly specify the type of the estimated application.

Regarding Question 4 we have observed that:

- None of the primary studies considered the particularities of each sourcing strategy.
- We did not identify any study considering offshore outsourcing as sourcing strategy. Therefore, there is a lack of research about effort estimation in offshore outsourced projects. Consequently, all the findings of this SLR are related to offshore insourced projects.
- The primary studies showed a trend in which there was always a site located in a country where salaries for software developers are quite high, and another site located in a country where salaries for developers are very low.
- The selected primary studies did not document the topology and the role of each site in the effort estimation process.

Finally, regarding Question 5, none of the selected studies explained how the data used to estimate the effort were extracted during the evaluated projects.

V. THREATS TO VALIDITY

In terms of threats to the validity of this SLR, the major issue is whether we have failed to find all the relevant primary studies. In order to mitigate that threat, we performed a very deep search strategy. We used keywords found in the most relevant works in effort estimation and Global Software Development.

We cannot claim that we retrieved all the available literature. However, we can state that we retrieved as many studies as possible, considering the restrictions that we have applied to our search string to reduce the number of irrelevant papers.

We would like to remark that one important task performed during the search process, the secondary search phase, was performed solely to the primary studies selected in the initial search phase, it was just a small portion of the complete set of studies retrieved in the search process. However, since the

inclusion and exclusion criteria of this SLR are very straight, we believe that most relevant works regarding the main topic of this work were evaluated.

It is important to note that only some of the initially retrieved studies complied with our inclusion criteria. Seven studies were excluded because they did not show any empirical evidence to support their proposals. Another three studies were excluded because of their quality. So, since we had just few studies to extract data and drawing conclusions about the state of the art of effort estimation in the GSD context, it is very hard to generalize the findings.

VI. CONCLUSIONS

This paper presented the results of a systematic literature review of effort estimation in the context of Global Software Development. The main goal of this paper was to present the state of the art regarding effort estimation in the context of GSD, in order to inform both research and practice.

The initial search phase returned 379 unique results and just 24 were selected. Out of those 24, 17 were excluded after a further investigation performed by reading the full text of those primary studies. The remaining 7 studies were then used in the secondary search phase, from which 1 more study was retrieved, bringing the total of selected studies to 8. Finally, after the achievement of the quality assessment process, we had a final list of 5 studies.

Only few studies complied with this SLR's inclusion criteria. Many studies were excluded because either they lack empirical evidence (eight) or have low quality (three).

It is important to note that none of the selected primary studies considered offshore outsourcing as sourcing strategy, it means that all the findings of this SLR are applicable to projects which apply offshore insourcing as sourcing strategy.

The identified effort estimation approaches are well-known in the context of co-located projects. The main difference appears to be in the used cost drivers. There is no standard effort estimation approach and the observed effort estimation approaches did not consider the uncertainty that is inherent to such domain. It was not possible to evaluate the accuracy of the observed approaches because the studies did not use the same accuracy metrics.

This SLR identified a wide range of cost drivers, where most of the primary studies presented cost drivers regarding cultural, language and time zone differences; these are factors directly related to a globally performed software project and they make the fulfillment of those projects more challenging.

The selected primary studies showed that the researchers are using just industrial single-company data; some studies did not document anything regarding the used data, as well as the type of the estimated applications.

Most of the questions regarding specific aspects of GSD projects were not addressed. Studies did not document the effects of the used sourcing strategy and the topology of the effort estimation process. On the other hand, we identified that the primary studies showed a trend in which there were always a site located in a developed country, where salaries

for software developers are quite high, and a site located in a developing country, where salaries for developers are very low. This was expected, since one of the biggest reasons to perform a software project globally is to reduce costs.

Finally, the selected primary studies did not state which activities of the development process were considered to calculate the effort estimation.

We believe that further investigation regarding effort estimation approaches used in co-located context is needed given the scarcity of primary studies identified in this SLR. We also put forward that the adaptation of those approaches considering the specific aspects of GSD and also the inherent uncertainty of the data [5] may provide more accurate effort estimates.

It seems that the GSD sourcing strategy and the effort estimation process topology can have a significant influence upon the effort estimates. For example, the transition of software or part of software developed by using the offshore outsourcing strategy seems to require more effort compared to the development carried out by using the offshore insourcing strategy. An effort estimation process performed in a distributed way seems to be more flexible and adequate to companies which apply agile methods.

However, since we did not find any evidence to support the above-mentioned assumptions, we believe that more investigation must be performed to confirm so. A single effort estimation technique may not be equally applicable to offshore outsourcing and offshore insourcing projects, thus warranting separate investigations that can later be combined. The same can be applied to the effort estimation process topologies.

Since we did not identified any study considering effort estimation in the context of offshore outsourced projects, we believe researches should explore this existent gap in order to provide solutions which could help practitioners in that kind of scenario.

Regarding the cost drivers, we believe that there is room for assessing their impacts on the effort and developing much more formal ways to measure them, which are directions for our future work.

VII. ACKNOWLEDGMENTS

We would like to thank FAPEMIG, CNPq, UFPI and INES, for partially supporting this work. We also would like to thank Professor Claes Wohlin for his feedback on a previous version of this paper.

REFERENCES

- [1] J. Herbsleb and D. J. Paulish, "Global software development at Siemens: Experience from nine projects," in *Proceedings of 27th International Conference on Software Engineering - ICSE'05*, St. Louis, USA, 2005, pp. 524–533.
- [2] J. Herbsleb, A. Mockus, T. A. Finholt, and R. E. Grinter, "Distance, dependencies, and delay in a global collaboration," in *Proceedings of the ACM Conference on Computer Supported Cooperative Work - CSCW'00*, New York, USA, 2000, pp. 319–328.
- [3] N. Fenton, W. Marsh, M. Neil, P. Cates, S. Forey, and M. Tailor, "Making resource decisions for software projects," in *Proceedings of 26th International Conference on Software Engineering - ICSE'04*, Edinburgh, Scotland, 2004, pp. 397–406.
- [4] E. Mendes, "Building a web effort estimation model through knowledge elicitation," in *Proceedings of 13th International Conference on Enterprise Information Systems - ICEIS'11*, Beijing, China, 2011, pp. 8–11.
- [5] —, "The use of bayesian networks for web effort estimation: Further investigation," in *Proceedings of 8th International Conference on Web Engineering - ICWE'08*, New York, USA, 2008, p. 203216.
- [6] —, "Using knowledge elicitation to improve web effort estimation: Lessons from six industrial case studies," in *Proceedings of 34th International Conference on Software Engineering - ICSE'12*, Zurich, Switzerland, 2012, pp. 1112–1121.
- [7] N. Ramasubbu and R. K. Balan, "Overcoming the challenges in cost estimation for distributed software projects," in *Proceedings of 34th International Conference on Software Engineering - ICSE'12*, Zurich, Switzerland, 2012, pp. 91–101.
- [8] N. C. Narendra, K. Ponnalagu, N. Zhou, and W. M. Gifford, "Towards a Formal Model for Optimal Task-Site Allocation and Effort Estimation in Global Software Development," in *Proceedings of 2012 Service Research and Innovation Institute Global Conference*, Silicon Valley, USA, 2012, pp. 470–477.
- [9] C. E. L. Peixoto, J. L. N. Audy, and R. Prikładnicki, "Effort Estimation in Global Software Development Projects: Preliminary Results from a Survey," in *Proceedings of 5th IEEE International Conference on Global Software Engineering - ICGSE'10*, Princeton, USA, 2010, pp. 123–127.
- [10] A. Lamersdorf, J. Munch, A. F. V. Torre, C. R. Sanchez, and D. Rombach, "Estimating the Effort Overhead in Global Software Development," in *Proceedings of 5th IEEE International Conference on Global Software Engineering - ICGSE'10*, Princeton, USA, 2010, pp. 267–276.
- [11] J. Kroll, J. L. N. Audy, and R. Prikładnicki, "Mapping the Evolution of Research on Global Software Engineering - A Systematic Literature Review," in *International Conference on Enterprise Information Systems - ICEIS 2011*, Beijing, China, 2011, pp. 260–265.
- [12] S. Schneider, R. Torkar, and T. Gorschek, "Solutions in global software engineering: A systematic literature review," *International Journal of Information Management*, vol. 33, no. 1, pp. 119–132, 2013.
- [13] S. Jalali and C. Wohlin, "Global software engineering and agile practices: A systematic review," *Journal of software: Evolution and Process*, vol. 24, no. 6, pp. 643–659, 2012.
- [14] E. Hossain, M. Ali Babar, and H. Y. Paik, "Using scrum in global software development: A systematic literature review," in *Proceedings of 4th IEEE International Conference on Global Software Engineering - ICGSE '09*, Limerick, Ireland, 2009, pp. 175–184.
- [15] F. Da Silva, C. Costa, A. França, and R. Prikładnicki, "Challenges and solutions in Distributed Software Development Project Management: A systematic literature review," in *Proceedings of 5th International Conference on Global Software Engineering - ICGSE 10*, Princeton, USA, 2010, pp. 87–96.
- [16] F. Q. B. Da Silva, R. Prikładnicki, A. C. C. Franca, C. V. F. Monteiro, C. Costa, and R. Rocha, "An evidence-based model of distributed software development project management: Results from a systematic mapping study," *Journal of software: Evolution and Process*, vol. 24, no. 6, pp. 625–642, 2012.
- [17] R. Prikładnicki and J. L. N. Audy, "Process models in the practice of distributed software development: A systematic review of the literature," *Information and Software Technology*, vol. 52, no. 8, pp. 779–791, 2010.
- [18] D. Mishra, A. Mishra, R. Colomo-Palacios, and C. Casado-Lumbreras, "Global software development and quality management: A systematic review," *Lecture Notes in Computer Science*, vol. 8186 LNCS, pp. 302–311, 2013.
- [19] D. Smite, C. Wohlin, T. Gorschek, and R. Feldt, "Empirical evidence in global software engineering: A systematic review," *Empirical Software Engineering*, vol. 15, no. 1, pp. 91–118, 2010.
- [20] I. Nurdiani, R. Jabangwe, D. Smite, and D. Damian, "Risk Identification and Risk Mitigation Instruments for Global Software Development: Systematic Review and Survey Results," in *Proceedings of 6th IEEE International Conference on Global Software Engineering Workshop - ICGSEW'11*, Helsinki, Finland, 2011, pp. 36–41.
- [21] M. J. Monasor, A. Vizcaino, M. Piattini, and I. Caballero, "Preparing students and engineers for global software development: A systematic review," in *Proceedings of 5th International Conference on Global Software Engineering - ICGSE 2010*, Princeton, USA, 2010, pp. 177–186.
- [22] S. Nidhra, M. Yanamadala, W. Afzal, and R. Torkar, "Knowledge transfer challenges and mitigation strategies in global software development-

- A systematic literature review and industrial validation," *International Journal of Information Management*, vol. 33, no. 2, pp. 333–355, 2013.
- [23] E. A. Matusse, E. H. M. Huzita, and T. F. C. Tait, "Metrics and indicators to assist in the distribution of process steps for distributed software development: A systematic review," in *Conference Proceedings of 38th Latin America Conference on Informatics - CLEI'12*, Medellin, Colombia, 2012, pp. 1–10.
- [24] I. Steinmacher, A. P. Chaves, and M. A. Gerosa, "Awareness support in distributed software development: A systematic review and mapping of the literature," *Computer Supported Cooperative Work*, vol. 22, no. 2-3, pp. 113–158, 2013.
- [25] V. Gomes and S. Marczak, "Problems? We All Know We Have Them. Do We Have Solutions Too? A Literature Review on Problems and Their Solutions in Global Software Development," in *Proceedings of IEEE 7th International Conference on Global Software Engineering - ICGSE'12*, Porto Alegre, Brazil, 2012, pp. 154–158.
- [26] N. D. Anh, D. S. Cruzes, and R. Conradi, "Dispersion, Coordination and Performance in Global Software Teams: A Systematic Review," in *Proceedings of the ACM-IEEE International Symposium on Empirical Software Engineering and Measurement (ESEM'12)*, Lund, Sweden, 2012, Proceedings Paper, pp. 129–138.
- [27] I. Steinmacher, A. P. Chaves, and M. A. Gerosa, "Awareness support in global software development: A systematic review based on the 3C collaboration model," *Lecture Notes in Computer Science*, vol. 6257 LNCS, pp. 185–201, 2010.
- [28] B. Kitcheham and S. Charters, "Guidelines for performing Systematic Literature Reviews in Software Engineering," Keele University, Tech. Rep., 2007.
- [29] M. Petticrew and H. Roberts, *Systematic Reviews in the Social Sciences: A Practical Guide*. Wiley-Blackwell, 2005.
- [30] H. Zhang and M. A. Babar, "On Searching Relevant Studies in Software Engineering," in *Proceedings of the 14th International Conference on Evaluation and Assessment in Software Engineering*, Keele University, England, 2010, pp. 1–10.
- [31] J. Herbsleb and D. Moitra, "Global software development," *IEEE Software*, vol. 18, no. 2, pp. 16–20, 2001.
- [32] P. Bjørndal, K. Smiley, and P. Mohapatra, "Global software project management: A case study," in *Lecture Notes in Business Information Processing*, vol. 54, no. SEAFOOD, Saint Petersburg, Russia, 2010, pp. 64–70.
- [33] T. Little, "Assessing the cost of Outsourcing: Efficiency, Effectiveness and Risk," in *Proceedings of IEEE International Conference on Exploring Quantifiable IT Yields - EQUITY '07.*, Washington DC, USA, 2007, pp. 95–100.
- [34] P. Keil, D. J. Paulish, and R. S. Sangwan, "Cost Estimation for Global Software Development," in *Proceedings of the 2006 International Workshop on Economics Driven Software Engineering Research*, New York, USA, 2006, pp. 7–10.
- [35] R. J. Madachy, "Cost modeling of distributed team processes for global development and software-intensive systems of systems," *Software Process Improvement and Practice*, vol. 13, no. 1, pp. 51–61, 2008.
- [36] M. Biehl and M. Torngren, "A cost-efficiency model for tool chains," in *Proceedings of IEEE Seventh International Conference on Global Software Engineering Workshops*, Porto Alegre, Brazil, 2012, pp. 6–11.
- [37] R. Madachy, "Distributed global development parametric cost modeling," *Lecture Notes in Computer Science*, vol. 4470 LNCS, pp. 159–168, 2007.
- [38] M. Muhairat, S. Aldaajeh, and R. E. Al-Qutaish, "The impact of global software development factors on effort estimation methods," *European Journal of Scientific Research*, vol. 46, no. 2, pp. 221–232, 2010.
- [39] A. C. M. Almeida, R. Souza, G. Aquino, and S. Meira, "Effort drivers estimation for brazilian geographically distributed software development," in *Lecture Notes in Business Information Processing*, vol. 35, no. SEAFOOD, Zurich, Switzerland, 2009, pp. 60–65.
- [40] J. Han and W. Jung, "Extracting communication structure of a development organization from a software repository," *Personal and Ubiquitous Computing*, pp. 1–9, 2013.
- [41] S. M. Handley and W. C. Benton Jr., "The influence of task- and location-specific complexity on the control and coordination costs in global outsourcing relationships," *Journal of Operations Management*, vol. 31, no. 3, pp. 109–128, 2013.
- [42] R. Kuni and N. Bhushan, "IT Application Assessment Model for Global Software Development," in *Proceedings of IEEE International Conference on Global Software Engineering - ICGSE'06*, Costao do Santinho, Brazil, 2006, pp. 92–100.
- [43] A. Lamersdorf and J. Münch, "Model-based task allocation in distributed software development," in *Lecture Notes in Business Information Processing*, vol. 54 LNBIP, no. SEAFOOD, Saint Petersburg, Russia, 2010, pp. 37–53.
- [44] S. Doma, L. Gottschalk, T. Uehara, and J. Liu, "Resource allocation optimization for GSD projects," in *Proceedings of International Conference Computational Science and Its Applications - ICCSA'09.*, vol. 2, Berlin, Germany, 2009, pp. 13–28.
- [45] M. Azzeh, "Software cost estimation based on use case points for global software development," in *Proceedings of 6th International Conference on Computer Science and Information Technology - ICCSIT'13*, Paris, France, 2013, pp. 214–218.
- [46] T. F. C. Tait and E. H. M. Huzita, "Software project management in distributed software development context," in *Proceedings of the 15th International Conference on Enterprise Information Systems - ICEIS'13*, vol. 2, Angers, France, 2013, pp. 216–222.
- [47] J. W. Kim and P. Meso, "Toward a framework of cultural effects on software development," in *Proceedings of 11th Americas Conference on Information Systems - AMCIS 2005*, vol. 6, Nebraska, USA, 2005, pp. 2718–2724.
- [48] M. M. Larsen, S. Manning, and T. Pedersen, "Uncovering the hidden costs of offshoring: The interplay of complexity, organizational design, and experience," *Strategic Management Journal*, vol. 34, no. 5, pp. 533–552, 2013.
- [49] K. Power, "Using silent grouping to size user stories," in *Lecture Notes in Business Information Processing*, vol. 77 LNBIP, no. XP2011, Madrid, Spain, 2011, pp. 60–72.
- [50] D. Smite and C. Gencel, "Why a CMMI level 5 company fails to meet the deadlines?" in *Lecture Notes in Business Information Processing*, vol. 32 LNBIP, no. PROFES, Oulu, Finland, 2009, pp. 87–95.
- [51] J. Dasgupta, "Monte carlo simulation based estimations: Case from a global outsourcing company," in *Proceedings of 2011 International Technology Management Conference - ITMC'11*, San Jose, USA, 2011, pp. 619–624.
- [52] S. Betz and J. Makio, "Amplification of the COCOMO II regarding Offshore Software Projects," in *Workshop on Offshoring of Software Development-Methods and Tools for Risk Management at the Second International Conference on Global Software Engineering*, Munich, Germany, 2007, pp. 1–11.
- [53] B. W. Boehm, *Software Engineering Economics*. Prentice Hall, 1981.
- [54] M. Cohn, *Agile Estimating and Planning*. Prentice Hall, 2005.
- [55] E. Mendes, *Cost Estimation Techniques for Web Projects*. IGI Publishing, 2007.
- [56] ISBSG. Isbsg database. [Online]. Available: <http://www.isbsg.org/>
- [57] P. Abrahamsson, R. Moser, W. Pedrycz, A. Sillitti, and G. Succi, "Software Estimation using the SLIM Tool," in *Proceedings of the Conference of the Centre for Advanced Studies on Collaborative Research*, 1992, pp. 439–475.
- [58] S. D. Conte, H. E. Dunsmore, and V. Y. Shen, *Software Engineering Metrics and Models*. Redwood City, CA, USA: Benjamin-Cummings Publishing Co., Inc., 1986.
- [59] A. Albrecht and J. E. Gaffney, "Software function, source lines of code, and development effort prediction: A software science validation," *Software Engineering, IEEE Transactions on*, vol. SE-9, no. 6, pp. 639–648, Nov 1983.
- [60] J. Schofield, A. Armentrout, and R. Trujillo, "Function points, use case points, story points: Observations from a case study," *CrossTalk*, vol. 26, no. 3, pp. 23–27, 2013.
- [61] J. L. Elshoff and M. Marcotty, "On the use of the cyclomatic number to measure program complexity," *SIGPLAN Not.*, vol. 13, no. 12, pp. 29–40, Dec. 1978. [Online]. Available: <http://doi.acm.org/10.1145/954587.954590>
- [62] D. Smite, C. Wohlin, Z. Galvina, and R. Prikladnicki, "An empirically based terminology and taxonomy for global software engineering," *Empirical Software Engineering*, vol. 19, no. 1, pp. 105–153, 2014.
- [63] R. Sangwan, M. Bass, N. Mullick, D. Paulish, and J. Kazmeier, *Global Software Development Handbook (Applied Software Engineering Series)*. Auerbach Publications, 2006.
- [64] S. L. Pfleeger, "Albert einstein and empirical software engineering," *Computer*, vol. 32, no. 10, pp. 32–38, Oct 1999.