On the Relationship Between Story Point and Development Effort in Agile Open-Source Software (Supplementary Material)

Vali Tawosi, Rebecca Moussa, Federica Sarro
{*vali.tawosi, rebecca.moussa.18, f.sarro*}*@ucl.ac.uk*
University College London, London, UK

Abstract

This document is supplementary to the paper entitled "On the Relationship Between Story Point and Development Effort in Agile Open-Source Software", accepted for presentation at the 16th ACM/IEEE International Symposium on Empirical Software Engineering and Measurement [1]. The data and scripts used in this study is available at:
`https://github.com/SOLAR-group/SPvsDevelopmentEffort.git`

Appendix

1 Data Selection

The data used in our study is sampled from the TAWOS dataset [2]. To sample data for RQ1 we used the following SQL query:

```
SELECT
   ID, Issue_Key, Resolution, Timespent, In_Progress_Minutes,
   Total_Effort_Minutes, Resolution_Time_Minutes, Story_Point
FROM
   Issue
WHERE
   Project_ID = <PROJECT_ID>
   AND Resolution IN ('Completed', 'Complete', 'Fixed', 'Done', 'Resolved', 'Implemented',
       'Answered', 'Handled by Support', 'Resolved Locally', 'Deployed', 'Community Answered')
   AND Timespent is not NULL;
```

To sample data for RQ2 and RQ3 we used the following SQL query:

```
SELECT
   ID, Issue_Key, Type, Status, Resolution, Creation_Date, Resolution_Date,
   Story_Point, Timespent, In_Progress_Minutes, Total_Effort_Minutes, Resolution_Time_Minutes
FROM
   Issue
WHERE
```

```
Project_ID = <PROJECT_ID>
AND Resolution IN ('Completed', 'Complete', 'Fixed', 'Done', 'Resolved', 'Implemented',
    'Answered', 'Handled by Support', 'Resolved Locally', 'Deployed', 'Community Answered')
AND Story_Point BETWEEN 1 AND 100;
```

2 Dataset Description

In total, our dataset has 37,440 unique issues from 37 projects coming from 13 different repositories:

Nine Projects from **Atlassian**:

Atlassian Bamboo (BAM) Bamboo is a continuous integration and continuous deployment server developed by Atlassian.

Atlassian Clover (CLOV) Clover is a Java code coverage analysis utility bought and further developed by Atlassian.

Atlassian Confluence Server (CONFSERVER) Confluence is a knowledge sharing tool that helps teams create and share content.

Atlassian Crowd (CWD) Atlassian Crowd is a Centralized identity management application.

Atlassian FishEye (FE) Atlassian FishEye is a revision-control browser and search engine.

Atlassian Jira Server, Software Cloud, ans Software Server (JRASERVER, JSWCLOUD, and JSWSERVER) Jira is a proprietary issue tracking product developed by Atlassian that allows bug tracking and agile project management.

Two Projects from **Apache**:

Apache Mesos (MESOS) is an open-source project to manage computer clusters. It was developed in C++ language at the University of California, Berkeley.

Apache Usergrid (USERGRID) is an open-source Backend-as-a-Service composed of an integrated distributed NoSQL database, application layer and client tier with SDKs for developers looking to rapidly build web and/or mobile applications. Apache Usergrid has been developed in Java.

Five Project from **Appcelerator**:

Aptana Studio (APSTUD) is an open-source integrated development environment (IDE) for building web applications.

Appcelerator DAEMON (DAEMON) The Appcelerator Daemon is a server that runs on a developer's computer and hosts services which power the tooling for Axway products such as Axway Titanium SDK.

Titanium Mobile Platform (TIDOC) Titanium Mobile is a mature platform for developers to build completely native cross-platform mobile applications.

The Titanium SDK (TIMOB) is the software development kit for Titanium platform.

Appcelerator Studio (TISTUD) Appcelerator Studio is an eclipse based IDE that provides a single, extensible environment to rapidly build, test, package, and publish mobile apps across multiple devices and OSs.

Five projects from **Hyperledger**:

Hyperledger Blockchain Explorer (BE) Hyperledger Explorer is a user-friendly Web application tool used to view, invoke, deploy or query blocks, transactions and associated data, network information (name, status, list of nodes), chain codes and transaction families, as well as any other relevant information stored in the ledger.

Hyperledger Fabric (FAB) Hyperledger Fabric is intended as a foundation for developing applications or solutions with a modular architecture. Hyperledger Fabric allows components,

such as consensus and membership services, to be plug-and-play. Its modular and versatile design satisfies a broad range of industry use cases. It offers a unique approach to consensus that enables performance at scale while preserving privacy.

Hyperledger Indy Node (INDY) Hyperledger Indy provides tools, libraries, and reusable components for providing digital identities rooted on blockchains or other distributed ledgers so that they are interoperable across administrative domains, applications, and any other silo. Indy is interoperable with other blockchains or can be used standalone powering the decentralization of identity.

Hyperledger Indy SDK (IS) Hyperledger Indy SDK provides a distributed-ledger-based foundation for self-sovereign identity. Indy provides a software ecosystem for private, secure, and powerful identity, and the Indy SDK enables clients for it. The major artifact of the SDK is a c-callable library; there are also convenience wrappers for various programming languages and Indy CLI tool.

Hyperledger Sawtooth (STL) Hyperledger Sawtooth offers a flexible and modular architecture separates the core system from the application domain.

One project from **Lsstcorp**:

Lsstcorp Data management (DM) Data Management is responsible for creating the software, services and systems which will be used to produce Rubin Observatory's data products.

One project from **Lyrasis**:

Lyrasis Dura Cloud (DURACLOUD) DuraCloud is a hosted service from LYRASIS that lets you control where and how your content is preserved in the cloud.

Four projects from **MongoDB**: MongoDB is a cross-platform document-oriented database program. Classified as a NoSQL database program, MongoDB uses JSON-like documents with schema.

MongoDB Compass (COMPASS) MongoDB Compass provides quick visualization of the structure of data in the database, and perform ad hoc queries – all with zero knowledge of MongoDB's query language.

The Mongo $C++$ driver (CXX) a $C++$ driver for MongoDB based on libmongoc.

The Mongo Java driver (JAVA) The official MongoDB Java Driver providing both synchronous and asynchronous interaction with MongoDB.

MongoDB Core Server (SERVER) MongoDB Enterprise Server is the commercial edition of MongoDB, available as part of the MongoDB Enterprise Advanced subscription.

One project from **Moodle**:

Moodle (MDL) Moodle is a free and open-source learning management system (LMS) written in PHP and distributed under the GNU General Public License.

Two project from **Mulesoft**:

Mule APIkit (APIKIT) APIkit is a tool for building Mule REST or SOAP APIs.

Mule (MULE) Mule is a lightweight enterprise service bus (ESB) and integration framework provided by MuleSoft. The platform is Java-based, but can broker interactions between other platforms such as .NET using web services or sockets.

One project from **Sonatype**:

Sonatype Nexus (NEXUS) Nexus is a repository manager. It allows developers to proxy, collect, and manage their dependencies.

One project from **Spring**:

Spring XD (XD) Spring XD is a unified, distributed, and extensible service for data ingestion, real time analytics, batch processing, and data export.

Four projects from **Talendforge**:

Talend Big Data (TBD) Talend Open Studio – Big Data is a free and open source tool for processing your data very easily on a big data environment.

Talend Data Preparation (TDP) Talend Data Preparation is a self-service application that enables information workers to prepare data for analysis and other data-driven tasks.

Talend Data Quality (TDQ) Talend's Data Quality profiles, cleans, and masks data in any format or size to deliver data you can trust for the insights you need.

Talend Enterprise Service Bus (TESB) Talend ESB is a reliable and scalable enterprise service bus (ESB) that lets development teams manage integration projects in a holistic manner, combining integration of applications and data management in complex and heterogeneous computing environments.

Talend Data Management (TMDM) Talend Data Management is a platform that combines all Talend products into a common set of solutions.

The descriptive statistics of the data used in the study are presented in the paper.

3 Additional Results

Table 1 shows the result of three correlation statistics for all three development time proxies and SP. The correlation of In-Progress Time with SP is discussed in the paper. Effort Time and Resolution Time show a relatively weaker correlation with SP, compared to In-Progress time. For Effort time in specific, the Spearman's ρ correlation is strong for four projects, medium for 15, and weak in 13 remaining cases. Regarding to Kendall's τ and Pearson r correlation coefficients, 50 cases out of 64 show a weak correlation while the remaining 14 cases have medium correlation. Finally, for Resolution time, out of 96 correlation coefficients only nine cases show a medium correlation and all 87 remaining cases show a weak correlaion.

Figure 1 and 2 show the plots of the regression fits, and Figure 3, 4, and 5 show the boxplots of the development time distribution for each story point value, for all projects analysed in RQ3.

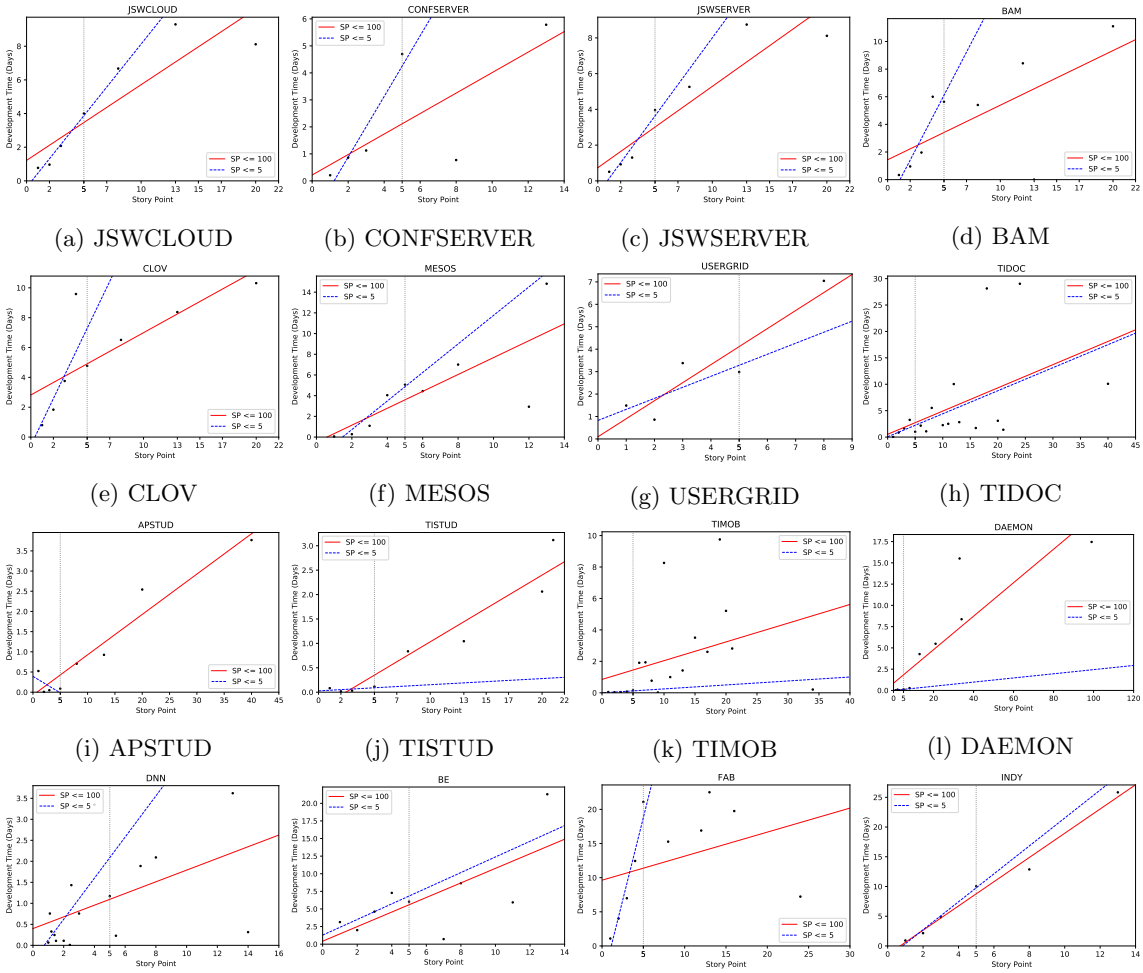
References

[1] V. Tawosi, R. Moussa, and F. Sarro, "On the relationship between story point and development effort in agile open-source software," in *16th ACM/IEEE International Symposium on Empirical Software Engineering and Measurement (ESEM'22)*. IEEE, 2022.

[2] V. Tawosi, A. Al-Subaihin, R. Moussa, and F. Sarro, "A versatile dataset of agile open source software projects," in *Proceedings of the 19th International Conference on Mining Software Repositories (MSR '22)*. IEEE, 2022.

Table 1: RQ2. Correlation results between SP and three development time proxies (p-value in brackets). Medium and strong correlations are highlighted in orange and red, respectively. Result of Shapiro-Wilk test on the distribution of In-Progress Time and SP values are also provided.

Project	In-Progress Time vs Story Point				Effort Time vs Story Point			Resolution Time vs Story Point		
	Shapiro Test	Spearman's ρ	Kendall's τ	Pearson r	Spearman's ρ	Kendall's τ	Pearson r	Spearman's ρ	Kendall's τ	Pearson r
JSWCLOUD	0.86 (<0.001)	0.54 (<0.001)	0.41 (<0.001)	0.47 (<0.001)	0.30 (<0.001)	0.23 (<0.001)	0.09 (0.238)	0.20 (0.006)	0.14 (0.008)	-0.05 (0.513)
CONFSERVER	0.76 (<0.001)	0.34 (<0.001)	0.26 (<0.001)	0.26 (<0.001)	0.10 (0.064)	0.07 (0.064)	-0.03 (0.571)	0.12 (0.020)	0.09 (0.023)	-0.02 (0.681)
JSWSERVER	0.86 (<0.001)	0.53 (<0.001)	0.40 (<0.001)	0.49 (<0.001)	0.29 (<0.001)	0.21 (<0.001)	0.09 (0.190)	0.21 (0.002)	0.15 (0.002)	0.02 (0.775)
BAM	0.80 (<0.001)	0.35 (<0.001)	0.28 (<0.001)	0.35 (<0.001)	0.21 (<0.001)	0.16 (<0.001)	0.12 (0.036)	0.00 (0.949)	0.00 (0.958)	-0.02 (0.692)
CLOV	0.81 (<0.001)	0.45 (<0.001)	0.34 (<0.001)	0.44 (<0.001)	0.43 (<0.001)	0.33 (<0.001)	0.39 (<0.001)	0.09 (0.289)	0.07 (0.269)	-0.03 (0.704)
MESOS	0.74 (<0.001)	0.40 (<0.001)	0.30 (<0.001)	0.35 (<0.001)	0.38 (<0.001)	0.29 (<0.001)	0.10 (<0.001)	0.16 (<0.001)	0.12 (<0.001)	0.04 (0.225)
USERGRID	0.86 (<0.001)	0.20 (0.013)	0.16 (0.008)	0.12 (0.139)	0.20 (0.013)	0.16 (0.008)	0.12 (0.139)	-0.04 (0.656)	-0.02 (0.722)	0.02 (0.827)
TIDOC	0.71 (<0.001)	0.48 (<0.001)	0.36 (<0.001)	0.27 (<0.001)	0.49 (<0.001)	0.36 (<0.001)	0.07 (0.078)	0.08 (0.034)	0.07 (0.021)	-0.07 (0.094)
APSTUD	0.75 (<0.001)	0.38 (<0.001)	0.30 (<0.001)	0.40 (<0.001)	0.38 (<0.001)	0.29 (<0.001)	0.40 (<0.001)	0.18 (0.002)	0.14 (0.001)	0.12 (0.042)
TISTUD	0.72 (<0.001)	0.42 (<0.001)	0.33 (<0.001)	0.35 (<0.001)	0.35 (<0.001)	0.27 (<0.001)	0.07 (0.005)	0.17 (<0.001)	0.13 (<0.001)	0.05 (0.036)
TIMOB	0.67 (<0.001)	0.28 (<0.001)	0.22 (<0.001)	0.23 (<0.001)	0.13 (<0.001)	0.10 (<0.001)	0.02 (0.419)	0.13 (<0.001)	0.10 (<0.001)	0.05 (0.019)
DAEMON	0.61 (<0.001)	0.62 (<0.001)	0.48 (<0.001)	0.66 (<0.001)	0.62 (<0.001)	0.47 (<0.001)	0.21 (0.018)	0.42 (<0.001)	0.31 (<0.001)	0.07 (0.435)
DNN	0.65 (<0.001)	0.32 (<0.001)	0.25 (<0.001)	0.27 (<0.001)	0.29 (<0.001)	0.22 (<0.001)	0.03 (0.264)	0.00 (0.927)	0.00 (0.964)	-0.06 (0.042)
BE	0.87 (<0.001)	0.19 (0.003)	0.15 (0.002)	0.21 (0.001)	0.29 (<0.001)	0.22 (<0.001)	0.11 (0.077)	0.28 (<0.001)	0.20 (<0.001)	0.08 (0.246)
FAB	0.79 (<0.001)	0.49 (<0.001)	0.38 (<0.001)	0.36 (<0.001)	0.50 (<0.001)	0.38 (<0.001)	0.21 (0.001)	0.49 (<0.001)	0.37 (<0.001)	0.20 (0.002)
INDY	0.89 (<0.001)	0.56 (<0.001)	0.44 (<0.001)	0.53 (<0.001)	0.45 (<0.001)	0.35 (<0.001)	0.28 (<0.001)	0.14 (0.004)	0.10 (0.004)	0.16 (0.001)
STL	0.89 (<0.001)	0.41 (<0.001)	0.32 (<0.001)	0.39 (<0.001)	0.40 (<0.001)	0.31 (<0.001)	0.08 (0.042)	0.34 (<0.001)	0.26 (<0.001)	0.12 (0.003)
IS	0.83 (<0.001)	0.49 (<0.001)	0.38 (<0.001)	0.43 (<0.001)	0.38 (<0.001)	0.28 (<0.001)	0.26 (<0.001)	0.25 (<0.001)	0.19 (<0.001)	0.10 (0.046)
DM	0.78 (<0.001)	0.49 (<0.001)	0.36 (<0.001)	0.42 (<0.001)	0.46 (<0.001)	0.34 (<0.001)	0.18 (<0.001)	0.42 (<0.001)	0.31 (<0.001)	0.17 (<0.001)
DURACLOUD	0.75 (<0.001)	0.52 (<0.001)	0.41 (<0.001)	0.47 (<0.001)	0.52 (<0.001)	0.41 (<0.001)	0.46 (<0.001)	0.29 (<0.001)	0.23 (<0.001)	0.09 (0.156)
COMPASS	0.78 (<0.001)	0.30 (<0.001)	0.23 (<0.001)	0.25 (<0.001)	0.24 (<0.001)	0.18 (<0.001)	0.21 (<0.001)	0.14 (0.024)	0.10 (0.023)	0.07 (0.223)
CXX	0.71 (<0.001)	0.27 (0.005)	0.22 (0.006)	0.36 (<0.001)	0.32 (0.001)	0.26 (0.001)	0.00 (0.998)	0.01 (0.906)	0.01 (0.906)	0.07 (0.457)
SERVER	0.73 (<0.001)	0.49 (<0.001)	0.37 (<0.001)	0.29 (<0.001)	0.51 (<0.001)	0.39 (<0.001)	0.15 (0.002)	0.43 (<0.001)	0.33 (<0.001)	0.06 (0.218)
EVG	0.76 (<0.001)	0.36 (<0.001)	0.28 (<0.001)	0.28 (<0.001)	0.34 (<0.001)	0.26 (<0.001)	0.07 (0.007)	0.28 (<0.001)	0.22 (<0.001)	0.11 (<0.001)
MULE	0.86 (<0.001)	0.48 (<0.001)	0.36 (<0.001)	0.49 (<0.001)	0.48 (<0.001)	0.36 (<0.001)	0.47 (<0.001)	0.10 (<0.001)	0.07 (<0.001)	-0.05 (0.035)
APIKIT	0.80 (<0.001)	0.37 (<0.001)	0.28 (<0.001)	0.30 (<0.001)	0.35 (<0.001)	0.26 (<0.001)	0.17 (0.005)	0.28 (<0.001)	0.21 (<0.001)	0.07 (0.266)
NEXUS	0.75 (<0.001)	0.25 (<0.001)	0.19 (<0.001)	0.25 (<0.001)	0.16 (0.001)	0.12 (0.001)	0.08 (0.083)	0.23 (<0.001)	0.18 (<0.001)	0.08 (0.115)
XD	0.80 (<0.001)	0.41 (<0.001)	0.31 (<0.001)	0.38 (<0.001)	0.36 (<0.001)	0.27 (<0.001)	0.09 (<0.001)	0.28 (<0.001)	0.21 (<0.001)	0.08 (0.003)
TDQ	0.84 (<0.001)	0.44 (<0.001)	0.32 (<0.001)	0.36 (<0.001)	0.27 (<0.001)	0.19 (<0.001)	0.13 (<0.001)	0.11 (<0.001)	0.08 (<0.001)	-0.04 (0.144)
TDP	0.85 (<0.001)	0.36 (<0.001)	0.27 (<0.001)	0.23 (<0.001)	0.25 (<0.001)	0.19 (<0.001)	0.13 (0.004)	0.16 (0.001)	0.12 (0.001)	0.06 (0.220)
TMDM	0.86 (<0.001)	0.18 (0.016)	0.14 (0.015)	0.23 (0.002)	0.16 (0.031)	0.12 (0.029)	0.09 (0.212)	-0.08 (0.267)	-0.06 (0.301)	-0.01 (0.848)
TESB	0.87 (<0.001)	0.33 (<0.001)	0.26 (<0.001)	0.24 (<0.001)	0.04 (0.507)	0.03 (0.510)	-0.05 (0.414)	0.14 (0.013)	0.11 (0.009)	0.00 (0.965)
Min		0.18	0.14	0.12	0.04	0.03	-0.05	-0.08	-0.06	-0.07
Max		0.62	0.48	0.66	0.62	0.47	0.47	0.49	0.37	0.20
Mean		0.40	0.31	0.35	0.33	0.25	0.15	0.19	0.14	0.05
SD		0.11	0.08	0.11	0.14	0.10	0.13	0.14	0.11	0.07



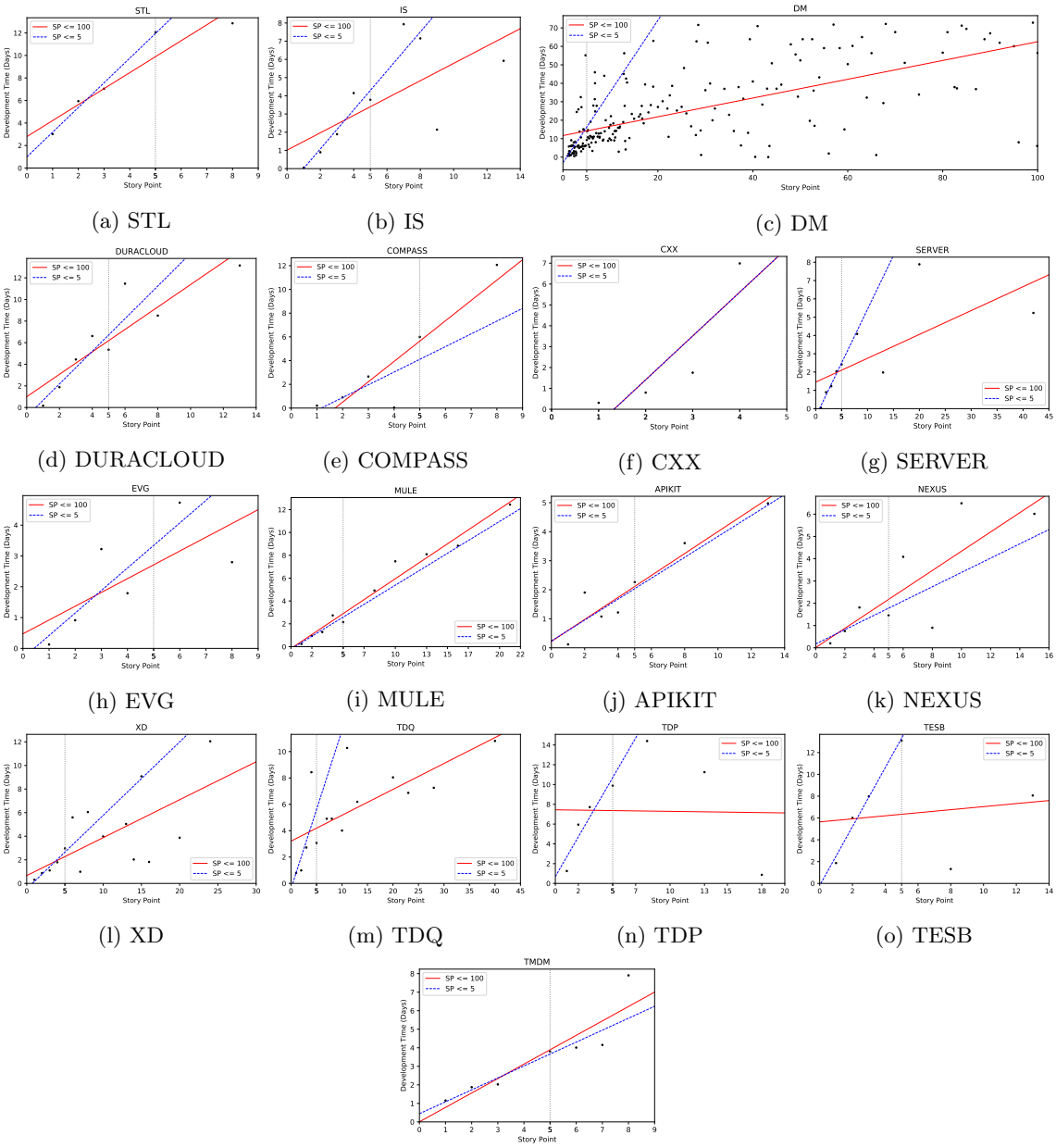
(a) JSWCLOUD (b) CONFSERVER (c) JSWSERVER (d) BAM

(e) CLOV (f) MESOS (g) USERGRID (h) TIDOC

(i) APSTUD (j) TISTUD (k) TIMOB (l) DAEMON

(m) DNN (n) BE (o) FAB (p) INDY

Figure 1: Linear Regression fit on the medians of the distributions of the development time (i.e., In-Progress time) values on each SP class value for projects, when SP less than or equal to 5 are considered (blue dashed Line) and SP less than 100 are considered (solid red line). The gray dotted vertical line marks the SP class 5.
Issues with In-Progress time <2 minutes and those with an outlier value are filtered out.



(a) STL (b) IS (c) DM

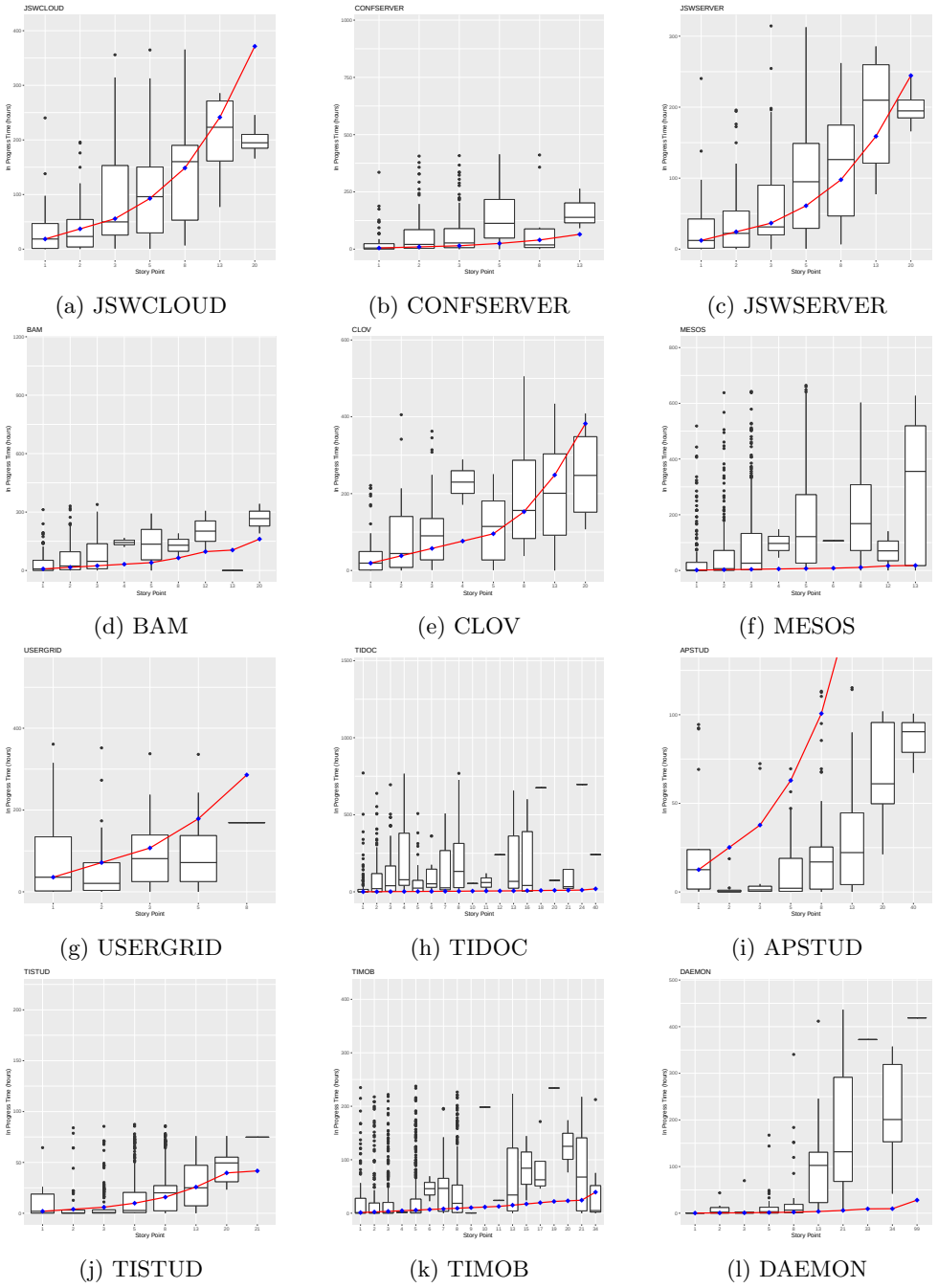
(d) DURACLOUD (e) COMPASS (f) CXX (g) SERVER

(h) EVG (i) MULE (j) APIKIT (k) NEXUS

(l) XD (m) TDQ (n) TDP (o) TESB

(p) TMDM

Figure 2: (Continued) Linear Regression fit on the medians of the distributions of the development time (i.e., In-Progress time) values on each SP class value for projects, when SP less than or equal to 5 are considered (blue dashed Line) and SP less than 100 are considered (solid red line). The gray dotted vertical line marks the SP class 5.
Issues with In-Progress time <2 minutes and those with an outlier value are filtered out.



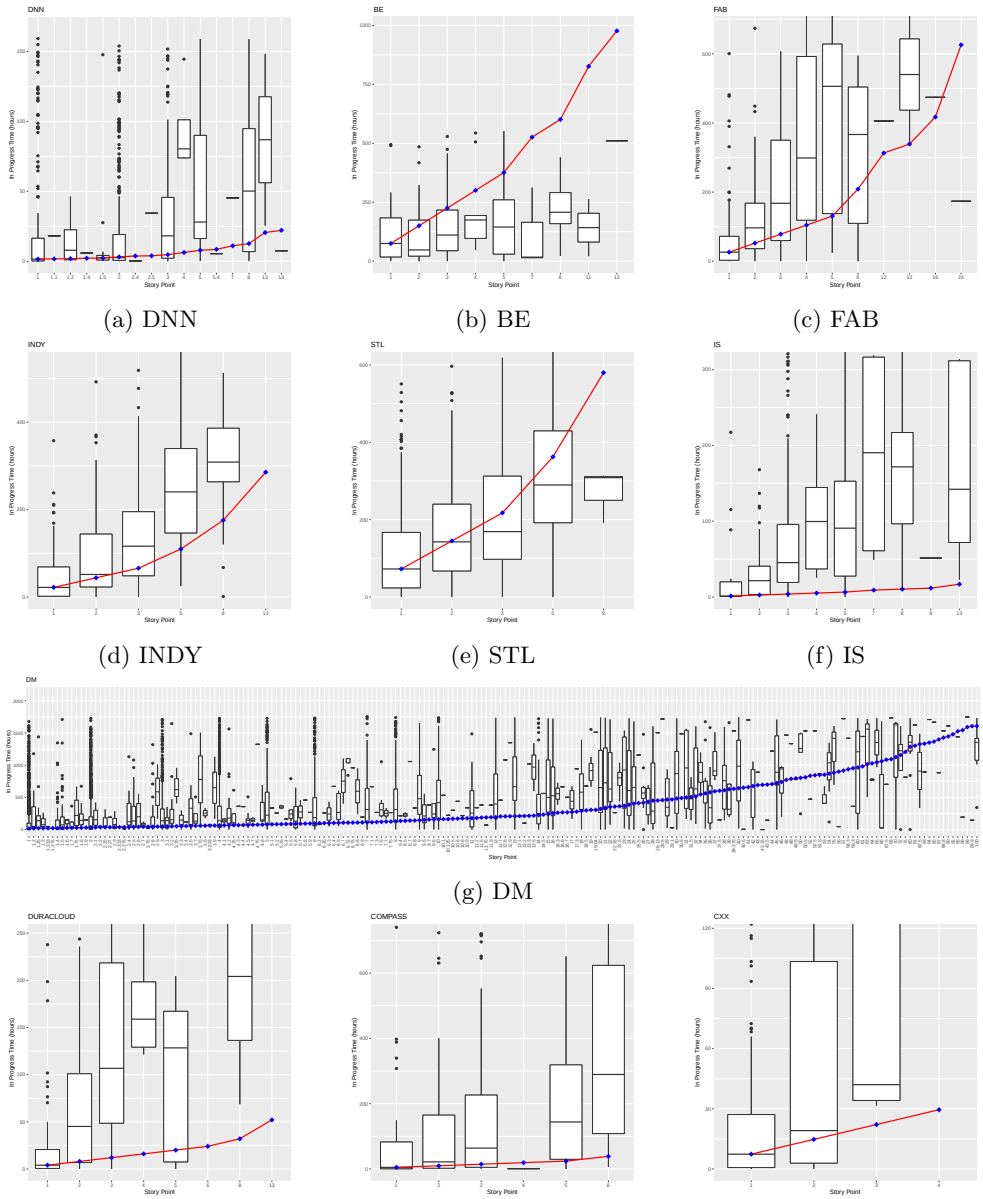
(a) JSWCLOUD (b) CONFSERVER (c) JSWSERVER

(d) BAM (e) CLOV (f) MESOS

(g) USERGRID (h) TIDOC (i) APSTUD

(j) TISTUD (k) TIMOB (l) DAEMON

Figure 3: Box plots of the distribution of development time per story point class. The red line depicts a project-specific baseline, drown based on the median development time for one story point.



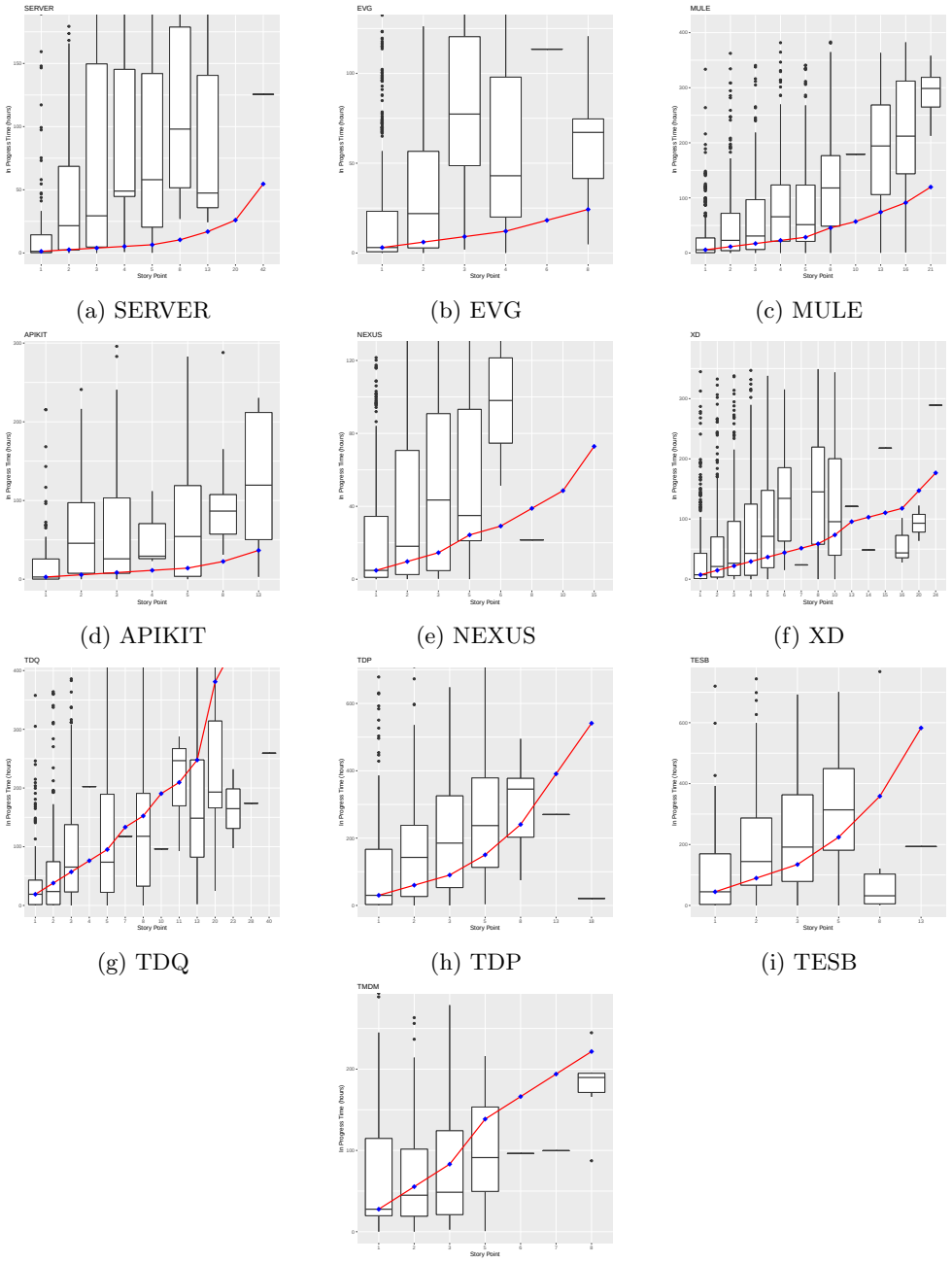
(a) DNN (b) BE (c) FAB

(d) INDY (e) STL (f) IS

(g) DM

(h) DURACLOUD (i) COMPASS (j) CXX

Figure 4: (Continued) Box plots of the distribution of development time per story point class. The red line depicts a project-specific baseline, drown based on the median development time for one story point.



(a) SERVER (b) EVG (c) MULE

(d) APIKIT (e) NEXUS (f) XD

(g) TDQ (h) TDP (i) TESB

(j) TMDM

Figure 5: (Continued) Box plots of the distribution of development time per story point class. The red line depicts a project-specific baseline, drown based on the median development time for one story point.