

# Forecasting in agilen Projekten mit der Monte-Carlo-Simulation

Herausgeber

**ESG Consulting GmbH**  
Livry-Gargan-Straße 6  
82256 Fürstenfeldbruck

Autor



Dr. Peter Merz  
[peter.merz@esg-consulting.com](mailto:peter.merz@esg-consulting.com)

Kontakte

Matthias Reimann  
Tel.: +49 (0)89 92161-2802  
E-Mail: [matthias.reimann@esg-consulting.com](mailto:matthias.reimann@esg-consulting.com)

Ulrich Bethäuser  
Tel.: +49 (0)89 92161-2517  
E-Mail: [ulrich.bethaeuser@esg-consulting.com](mailto:ulrich.bethaeuser@esg-consulting.com)

## Inhalt

<b>1. Voraussetzung für ein realistisches Forecasting</b> .....	<b>5</b>
<b>2. Analyse historischer Daten</b> .....	<b>7</b>
<b>3. Forecasting mit der Monte-Carlo-Simulation</b> .....	<b>9</b>
<b>4. Probabilistisches Forecasting mit Epics eines Groß-Projektes</b> .....	<b>10</b>
<b>5. Visualisierung der Prozess-Stabilität mit dem „Cumulative Flow Diagram (CFD)“</b> .....	<b>17</b>
<b>6. Schätzen von User-Stories ist nicht zielführend</b> .....	<b>19</b>
<b>7. Epilog</b> .....	<b>21</b>
<b>8. Abkürzungen</b> .....	<b>23</b>

## Abbildungsverzeichnis

Abbildung 1: einfaches Queuing System .....	5
Abbildung 2: Ausgangsdaten mit Auswahlmöglichkeit.....	7
Abbildung 3: Auswertung der Einzel-Items (historische Daten) .....	8
Abbildung 4: Simulations-Ergebnisse für den Forecast .....	10
Abbildung 5: Auswertung aller Epics ab Sprint 70 mit Ausgabe in Sprint-Einheiten .....	11
Abbildung 6: Auswertung aller Epics aus Jira mit Ausgabe in Tages-Einheiten .....	11
Abbildung 7: simulierter Forecast für das Großprojekt.....	12
Abbildung 8: Entwicklung des Durchsatzes über die Zeit .....	12
Abbildung 9: Entwicklung des Durchsatzes seit Anfang 2022 .....	13
Abbildung 10: Aging-Chart: aktuelle Dauer der Epics in Arbeit.....	13
Abbildung 11: Auswertung aller Epics mit Enddatum $\geq$ 01.10.2021 .....	14
Abbildung 12: simulierter Forecast mit Epics, die ab September 2021 fertiggestellt wurden.....	14
Abbildung 13: Simulation einer möglichen Anzahl Fertigstellungen von Epics .....	15
Abbildung 14: Auswertungen, nachdem Epics mit einer Cycletime $\geq$ 300 herausgefiltert wurden.....	16
Abbildung 15: CFD des Großprojektes ab Umstieg auf Jira .....	17
Abbildung 16: Zoom in das CFD .....	19
Abbildung 17: Dauer von User-Stories mit unterschiedlichen Story-Points .....	20

Abbildung 18: Zusammenfassung der Auswertung Story-Point vs. Dauer von User-Stories über alle 4 Teams ... 21  
Abbildung 19: Numerische Auswertung Story-Points versus Dauer von User-Stories ..... 21

## Tabellenverzeichnis

Tabelle 1: Gegenüberstellung der Messwert mit gefilterten Epics mit Epic-Fertigstellung ab 01.09.2022 ..... 15  
Tabelle 2: Gegenüberstellung der Messwert mit gefilterten Epics anhand Cycletime ..... 16  
Tabelle 3: Abkürzungen ..... 23

## Zusammenfassung

Die meisten Projekte im agilen Umfeld nutzen Story-Points zum Schätzen von User-Stories; zur Sprint-Planung wird hierzu die Anzahl Story-Points herangezogen, welche ein Team durchschnittlich pro Sprint absolvieren kann (= Plan-Velocity). Das Forecasting in Projekten basiert oftmals auf der Plan-Velocity der Teams oder, wenn es realistischer wirken soll, auf den Velocity-Durchschnittswerten aus der Vergangenheit. Dies kann hinsichtlich mehrerer Faktoren suboptimal sein:

- Wenn das Management nach realistischen Fertigstellungsterminen fragt, wird häufig mit der für das Management unverständlichen Währung der Velocity und Story-Points argumentiert.
- Für den Forecast werden normalverteilte Velocities angenommen, um somit einen linearen Bereich von möglichen Fertigstellungsterminen zu errechnen. Velocities entsprechen jedoch bei weitem nicht den Dauern von Arbeitspaketen (Work-Items) und die Dauer der Work-Items ist bei weitem nicht normalverteilt. Somit sind solche Forecasts mit großer Skepsis zu betrachten.
- Wenn die Prozess-Stabilität nicht genau beobachtet und justiert wird, können die Schätzungen mit Story-Points und die Planung mit Velocities ein stark verzerrtes Bild des Projekt-Status abgeben.

Um ein realistisches und probabilistisches Forecasting zu etablieren, muss zum einen der Umsetzungs-Prozess mit geeigneten Mitteln genauestens beobachtet und justiert werden und die zugrunde liegenden Daten müssen hinsichtlich Cycle-Time probabilistisch z.B. via Monte-Carlo-Simulation (MCS) ausgewertet werden. Die ESG-Consulting GmbH hat hierfür ein Werkzeug (Excel mit VBA) erstellt. Hiermit können folgende Analysen sehr einfach durchgeführt werden:

- Analyse von sog. Single-Items, seien dies User-Stories, Epics, Themes etc. Mittels Scatterplots und Histogrammen können Auswertungen zur Bearbeitungsdauer der Work-Items mit Angabe von Perzentilen angestellt werden.
- Der Durchsatz an Work-Items je Zeiteinheit (Tag, Woche, Monat, Sprint) kann über die Projektlaufzeit grafisch ausgewertet werden.
- Die aktuell in Bearbeitung befindlichen Work-Items können mit dem Aging-Chart hinsichtlich ihrer Bearbeitungsdauer grafisch ausgewertet und in Relation zu den gewählten Perzentilen gesetzt werden, um somit rasch auf sich hinauszögernde Work-Items reagieren und Maßnahmen ergreifen zu können.
- Mit dem Cumulative-Flow-Diagramm (CFD), welches dynamisch konfiguriert werden kann, wird man in die Lage versetzt, eine Auswertung über „Ankunftsrate“ und „Abflugrate“, über WIP, Durchsatz und approximierter Cycle-Time je gewähltem Zeitintervall durchzuführen.
- Weiterhin kann mit einer integrierten Monte-Carlo-Simulation eine Prognose über die Wahrscheinlichkeit von Fertigstellungsterminen bzw. voraussichtlich bis zu einem gewünschten Termin fertigstellbarer Work-Items inklusive Perzentile erstellt werden.

Die Angabe von geplanten Story-Points und von Velocity ist für all diese Auswertungen nicht nötig. Als Input benötigt man lediglich die Work-Items mit jeweiligem Datum, wann diese in Bearbeitung gingen und Enddatum bei den Work-Items, die bereits geschlossen wurden. Diese Daten können z.B. aus Jira importiert werden.

## 1. Voraussetzung für ein realistisches Forecasting

Forecasts zu erstellen, ohne eine Vorstellung von der System- bzw. - Prozess-Stabilität zu haben, ist kaum besser als bloße Vermutungen. Um ein realistisches Forecasting zu betreiben, muss der zugrunde liegende Prozess bzw. die Vorgehensweise probabilistisch vorhersagbar und entsprechend stabil sein. Zur Prozess-Stabilität gibt es in der Queuing-Theorie das sog. „Little’s Law“ (1961 von Dr. John D. C. Little definiert). Diesem Gesetz liegt folgende Beobachtung zu Grunde, welches auf beliebige Systeme angewandt werden kann:

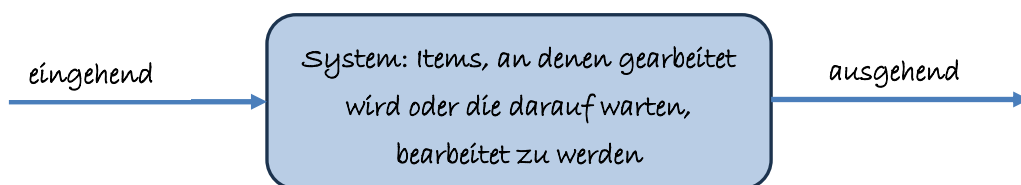


Abbildung 1: einfaches Queuing System

Das „Little’s Law“ ist definiert:

$$L = \lambda * W$$

wobei gilt:

**L** = durchschnittliche Anzahl Items im Queuing-System

**λ** = durchschnittliche Anzahl Items, die per Zeiteinheit ankommen

**W** = durchschnittliche Wartezeit eines Items im System

Dieses Gesetz sagt nicht, wie Manager Nachteile oder Defizite im System behandelt sollen, es zeigt aber eine nötige Relation zwischen den Messungen (L, λ und W).

In den späten 80er Jahren nutzte die Operations Management Community (OM) dieses Gesetz, um es an deren Bedürfnisse bzw. Perspektiven anzupassen:

$$\text{Durchschnittliche Cycletime} = \frac{\text{Durchschnittlicher Work in Progress (WIP)}}{\text{Durchschnittlicher Durchsatz}}$$

**Cycle Time:** Die durchschnittliche Zeit, die ein Item benötigt, um durch ein System / einen Prozess zu fließen

**Work In Progress (WIP):** durchschnittliche Anzahl nicht geschlossener und bereits begonnener Items im System, die also „In Progress“ sind.

**Durchsatz:** der durchschnittliche Durchsatz im System, die Menge abgeschlossener Arbeit bzw. Items pro Zeiteinheit

Gleichung 1 (Little's Law original) bezieht sich ausdrücklich auf die Ankunftsrate, wohingegen Gleichung 2 sich ausdrücklich auf die „Abflugrate“ (die fertiggestellten Items) bezieht. Beide besagen aber, je mehr Items, an denen in irgendeiner Zeitdauer gearbeitet wird, im System sind, desto länger dauert es, bis solche Items durchschnittlich abgeschlossen werden.

Little's Law ist der wichtigste Grund für die Umstellung auf einen stärker WIP-begrenzten, Pull-basierten Prozess mit kontinuierlichem Fluss!

*👉 Die Größe der Items spielt keine Rolle, da Little's Law auf Beziehungen zwischen Durchschnittswerten basiert. Dies bedeutet: für ein Forecasting, basierend auf einem stabilen System / Prozess, werden keine Schätzungen wie Story-Points benötigt.*

Das größere Forecasting-Problem wird normalerweise durch einen zu großen WIP bedingt oder aber auch durch die Häufigkeit, mit welcher die Annahmen von Little's Law verletzt werden.

Zusammengefasst, was Little's Law besagt ist: wenn der WIP im Durchschnitt vergrößert wird, dann resultiert daraus eine Veränderung der Cycletime und oder des Durchsatzes.

Um einen Prozess zu verbessern, wäre es auch hilfreich, die Flow-Efficiency zu betrachten:

$$\text{Flow Efficiency} = \frac{\text{Cycletime} - \text{Wartezeit} - \text{blockierte Zeit}}{\text{Cycletime}}$$

Wie bereits erwähnt: deterministische Forecasts stellen keine seriöse Grundlage für Entscheidungen dar. Viel aussagekräftiger sind Forecasts, die Projektionen in die Zukunft zum einen mit Wahrscheinlichkeiten des Eintretens und mit Bandbreiten von Termin-Dimensionen berechnet und dargestellt werden. Hierfür sind die Voraussetzungen:

- Ein stabiler Prozess
- Eine Auswahl von repräsentativen historischen Einzeldaten in Form von Cycletime und Abschluss-Zeitpunkt von Items (z.B. Epics)
- Stetige Auswertung mit Hilfe von neu ankommenden Daten. Wettervorhersagen zum Beispiel werden auch nicht nur einmal pro Woche berechnet, sondern die sich ändernden Parameter aus den Messstationen werden permanent in das Vorhersage-Modell aufgenommen, um somit zu einer besseren unmittelbaren Vorhersage zu kommen.
- Die ersten erhobenen Messwerte werden auf Grund eines noch nicht stabilen Systems noch erhebliche Ungenauigkeiten aufweisen. Der Forecast kann eben nur durch stetige Verbesserung der Systemstabilität und durch Hinzufügen neuer Messwerte verbessert werden.

## 2. Analyse historischer Daten

Cycletime ist die Fluss-Metrik, die benutzt wird, um Forecasts für „Single Items“ zu erstellen. Um einen realistischen Gesamt-Forecast z.B. eines Projektes machen zu können, muss zuerst analysiert werden, ob es sich überhaupt um ein stabiles System (bzw. Projekt-Prozess) handelt. Hierfür müssen für die zu messenden Items (z.B. Epics) lediglich die Cycletimes und das Datum bzw. der Sprint, wann das Item geschlossen wurde, gesammelt werden. Jetzt besteht aber auch die Kunst darin, den Bereich von vorhandenen Daten auszuwählen, der für eine Prognose überhaupt herangezogen werden kann. Denn oftmals gibt es in Projekten bestimmte Zeiten, in welchen höhere Wartezeiten bestehen (z.B. Urlaubszeit, Feiertage, ungeplante Ausfallszeiten von Infrastruktur, hohe Variabilität von externen Lieferungen).

In dem Tool können die Input-Daten in das Dashboard kopiert (oder auch automatisiert importiert) werden:

	A	B	C	D	E
1	Start	Ende	Cycletime		
2	Datum	Datum	Tage	Daten auswählen	
3		24.04.2023	20	Erste Zeile	5
4		05.05.2023	15	Letzte Zeile	18
5		10.05.2023	23		
6		18.05.2023	8		
7		23.05.2023	17		
8		31.05.2023	28		
9		12.06.2023	9		
10		13.06.2023	13		
11		21.06.2023	24		
12		29.06.2023	31		
13		12.07.2023	11		
14		13.07.2023	15		
15		18.07.2023	8		
16		19.07.2023	20		
17		26.07.2023	15		
18		26.07.2023	16		

Abbildung 2: Ausgangsdaten mit Auswahlmöglichkeit

Die über die Zellen E4 und E5 ausgewählten Zeilen für den zu benutzenden Input-Bereich werden grün markiert. Zu beachten wäre, dass man entweder die Cycletime aus Start- und Ende-Datum berechnen lässt oder aber die Cycletimes aus dem führenden System (z.B. Jira) zusammen mit Fertigstellungs-Datum importiert. Statt Datum kann auch „Sprint“ (Zelle B2 ist ein Auswahlfeld) als Zeiteinheit benutzt werden. Die anschließenden Berechnungen und Grafiken werden entsprechend angepasst.

Als nächsten Schritt können die Auswertungen der Einzel-Items angestoßen werden (siehe Abbildung 1). Hierbei werden die Cycletimes sowohl als sog. Scatterplot und als Histogramm ausgegeben.

Mit dem Scatterplot werden die Cycletimes der Items (Y-Achse) entsprechend deren jeweiligen Abschluss-Datums (X-Achse) ausgegeben. Mit dem Histogramm werden die Anzahl Items je Cycletime ausgegeben.

In der Grafik grün markierte Felder sind immer Eingabe-Felder, rot markierte Felder sind berechnete Felder.

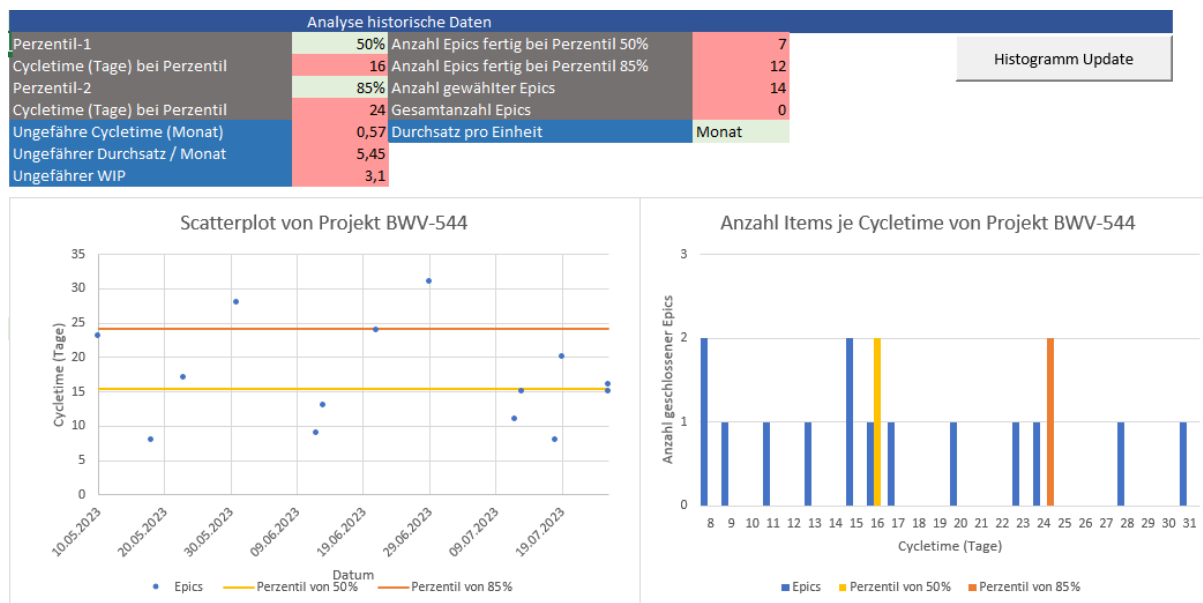


Abbildung 3: Auswertung der Einzel-Items (historische Daten)

Ein wichtiger Aspekt beim Forecasting ist die Angabe von **Perzentilen**, denn valide Prognosen liegen in Form einer Spanne von Ergebnissen und einer mit dieser Spanne verbundenen Wahrscheinlichkeit vor. Die Perzentile-Linien im obigen Scatterplot und im Histogramm entsprechen genau dieser Definition.

- Perzentil-1: hier 50%, bedeutet,
  - » dass in diesem Szenario 50% der Items eine Cycletime von 16 oder weniger Tagen haben
  - » bei einer Anzahl von insgesamt 14 ausgewählten Items wurden 7 Items (mit einer Cycletime bis zu 16 Tagen) fertig.
- Perzentil-2: hier 85%, bedeutet,
  - » dass in diesem Szenario 85% der Items eine Cycletime von 24 oder weniger Tagen haben.
  - » bei einer Anzahl von insgesamt 14 ausgewählten Items wurden 12 Items (mit einer Cycletime bis zu 24 Tagen) fertig.
- Für die Berechnungen des WIPs kann man in Zelle K8 die Zeiteinheit vorgeben. Wurde in Zelle B2 statt „Datum“ der Wert „Sprint“ angegeben, so steht auch hier nur die Zeiteinheit Sprint zur Verfügung. Ansonsten kann man hier zwischen den Zeiteinheiten Tag, Woche und Monat wählen. In diesem Szenario wurde der Wert Woche gewählt.
- Prozess-Qualitäten anhand Little's Law:
  - » Ungefähre Cycletime (Woche): die zur Auswertung herangezogenen Items hatten eine durchschnittliche Cycletime von 2,43 Wochen



- » Ungefährer Durchsatz / Woche: im Durchschnitt wurden 1,27 Items pro Woche abgeschlossen
- » Ungefährer WIP: dieser Wert wird aus den beiden oberen berechnet und ergibt einen in Bearbeitung befindlichen Bestand von durchschnittlich 3,1 Items pro Woche. Hinweis: wenn man die Zeiteinheit auf z.B. Monat ändert, dann verändert sich natürlich der durchschnittliche WIP nicht, da die beiden anderen Werte dieselbe Zeiteinheit benutzen. Man darf nie Messwerte mit unterschiedlichen Zeit-Einheiten nutzen!

### 3. Forecasting mit der Monte-Carlo-Simulation

Vorausgesetzt, es existieren valide Messergebnisse zu den Items, kann auch eine Simulation mit Hilfe der Monte-Carlo-Methode durchgeführt werden. Abbildung 4 zeigt das Ergebnis für die Input-Daten, wie sie in Abbildung 2 zur Verfügung gestellt wurden.

Für die MCS wird der Durchsatz benötigt. Dieser wird aus der Cycletime der Items berechnet, es werden also keine weiteren Daten bzw. Messwerte wie z.B. Story-Points für die Simulation benötigt!

Die Simulation startet mit dem Aufbau eines Arrays von Zeitstempel zwischen dem ersten und dem letzten Abschluss-Datum der Items und zählt die Anzahl der beim jeweiligen Zeitstempel fertig gestellten Items. Anhand dieses Arrays wird die Simulation durchgeführt und die Ergebnisse für die Aufbereitung der beiden Simulations-Grafiken verwendet. Mit der MCS werden in einer Schleife mit  $n$  Schritten ( $n$  = Anzahl Simulations-Schritte) so lange per Zufallszahl die Anzahl abgeschlossener Items je Zeitscheibe aus dem Input-Array generiert, bis die Anzahl von noch nicht abgeschlossenen Items erreicht ist. Somit erhält man eine stochastisch unabhängige Verteilung von Anzahlen Zeiteinheiten, bis alle noch nicht abgeschlossenen Items abgeschlossen werden könnten. Die linke Grafik in Abbildung 4 stellt die sog. Verteilungsfunktion dar, d.h. die Häufigkeit zu einem bestimmten Zeitpunkt simulierter geschlossener Items.

Die rechte Grafik ist die kumulierte Verteilungsfunktion, d.h. die Anzahl der geschlossenen Items wird mit fortschreitendem Zeitstempel aufsummiert. Die Anzahl Items ganz rechts muss immer die Anzahl der Simulations-Schritte ergeben.

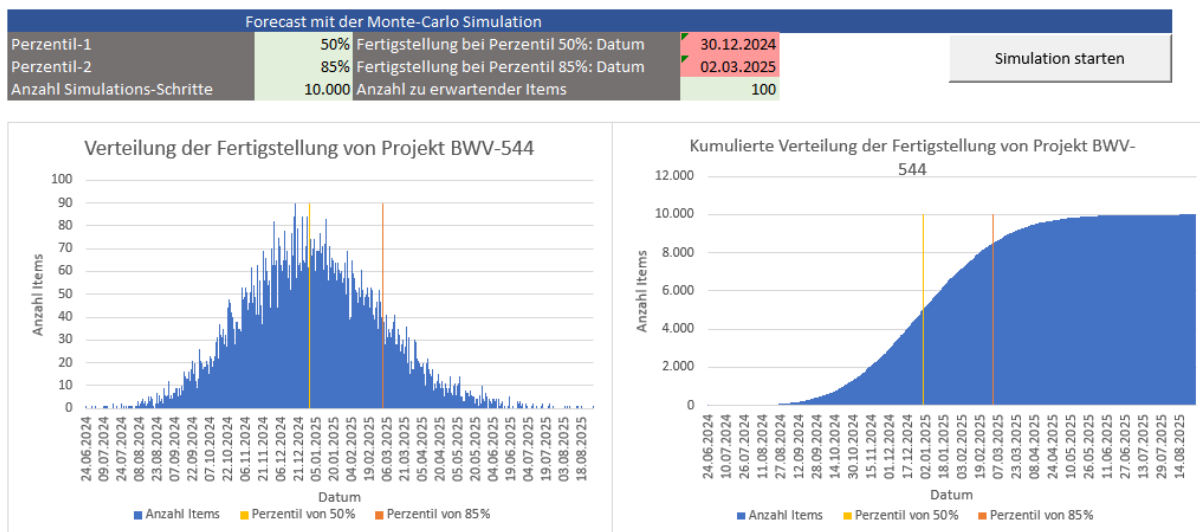


Abbildung 4: Simulations-Ergebnisse für den Forecast

Auch hier können wieder Parameter definiert werden:

- Anzahl zu erwartender Items: dies ist die Anzahl an Items (z.B. Epics), die noch im Backlog und nicht abgeschlossen sind. In diesem Falle wird von 100 noch nicht abgeschlossenen Items ausgegangen. Dieser Wert geht in die MCS ein.
- Anzahl Simulations-Schritte: gibt an, wie viele Permutationen die MCS durchführen soll. Je höher die Anzahl an Permutationen, desto ausgeglichener wird die Verteilungsfunktion. 10.000 Schritte sind völlig ausreichend.
- Perzentil-1: hier 50%: bedeutet, mit einer Wahrscheinlichkeit von 50% können die noch offenen Items bis zum 30.12.2024 abgeschlossen werden.
- Perzentil-2: hier 85%: bedeutet, mit einer Wahrscheinlichkeit von 85% können die noch offenen Items bis zum 02.03.2025 abgeschlossen werden.

## 4. Probabilistisches Forecasting mit Epics eines Groß-Projektes

Im Folgenden sollen die Daten, die mittels Jira in einem agilen Groß-Projekt gesammelt wurden, ausgewertet werden. Das Projekt wurde nach Scrum Nexus durchgeführt und ein Sprint hatte eine Dauer von 2 Wochen. Vier Entwicklungsteams, ein Fachteam und ein Integrationsteam waren für die Umsetzung verantwortlich. Das Integrationsteam war hauptsächlich für DevOps und die Service-Architektur zuständig. Es sei zu bemerken, dass das Projekt erst ab Sprint 70 vollumfänglich auf JIRA umgestiegen ist. Daher beginnt die Auswertung für den Forecast hier erst mit Sprint 70.

Analyse historische Daten			
Perzentil-1	50%	Anzahl Epics fertig bei Perzentil 50%	11
Cycletime (Sprint) bei Perzentil	5	Anzahl Epics fertig bei Perzentil 85%	37
Perzentil-2	85%	Anzahl gewählter Epics	224
Cycletime (Sprint) bei Perzentil	17	Gesamtanzahl Epics	224
Ungefähre Cycletime (Sprint)	8,19	Durchsatz pro Einheit	Sprint
Ungefährer Durchsatz / Sprint	3,92		14
Ungefährer WIP	32,2	Anzeige bis Maximum Cycletime (Tage)	0

Historische Daten Update	
Scatterplot-Zeitraum	
Abgeschlossen seit	02.03.2021
Abgeschlossen bis	10.05.2023

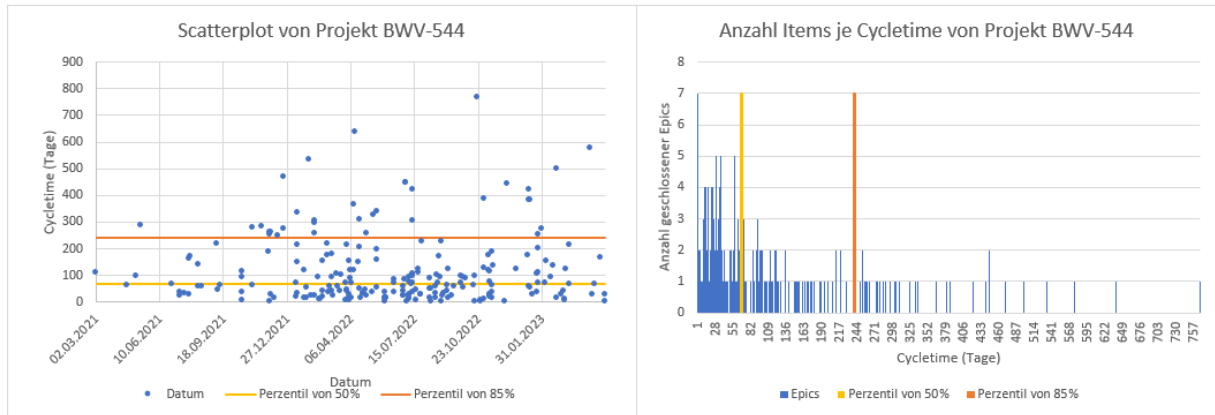


Abbildung 5: Auswertung aller Epics ab Sprint 70 mit Ausgabe in Sprint-Einheiten

Analyse historische Daten			
Perzentil-1	50%	Anzahl Epics fertig bei Perzentil 50%	112
Cycletime (Tag) bei Perzentil	68	Anzahl Epics fertig bei Perzentil 85%	190
Perzentil-2	85%	Anzahl gewählter Epics	224
Cycletime (Tag) bei Perzentil	240	Gesamtanzahl Epics	224
Ungefähre Cycletime (Tag)	114,70	Durchsatz pro Einheit	Tag
Ungefährer Durchsatz / Tag	0,28		14
Ungefährer WIP	32,2	Anzeige bis Maximum Cycletime (Tage)	0

Historische Daten Update	
Scatterplot-Zeitraum	
Abgeschlossen seit	02.03.2021
Abgeschlossen bis	10.05.2023

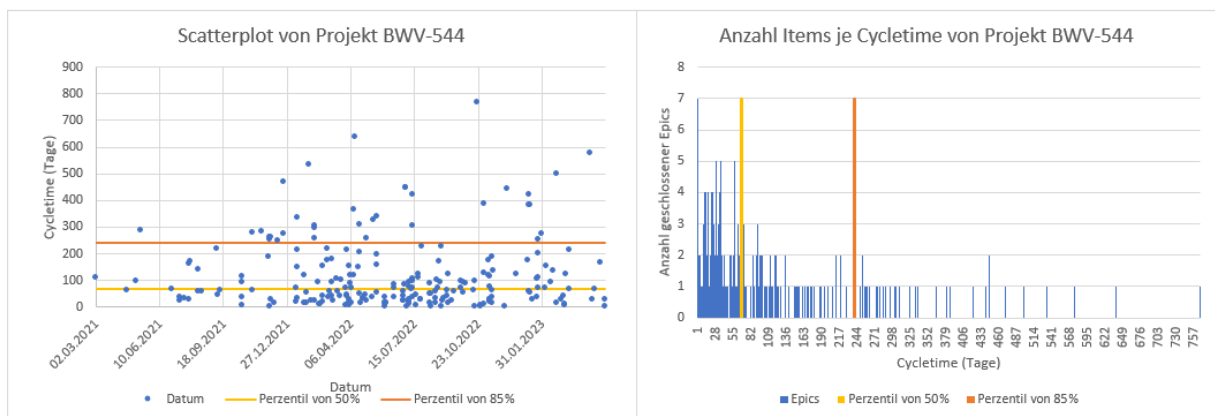


Abbildung 6: Auswertung aller Epics aus Jira mit Ausgabe in Tages-Einheiten

Im Zeitraum vom 02.03.2021 bis 10.05.2023 wurden 224 Epics abgearbeitet. Im Backlog befanden sich noch 145 Epics. Von den 224 geschlossenen Epics haben 85% eine Cycletime <= 240 Tagen und 50% haben eine Cycletime von <= 68 Tagen. Der durchschnittliche Durchsatz an Epics pro Woche liegt bei 1,96 Epics und es befanden sich im Durchschnitt 16,4 Epics parallel in Bearbeitung (WIP). Bei 4 Entwickler-Teams ist dieser WIP als zu hoch anzusehen, was sich auch bei der langen durchschnittlichen Cycletime widerspiegelt.

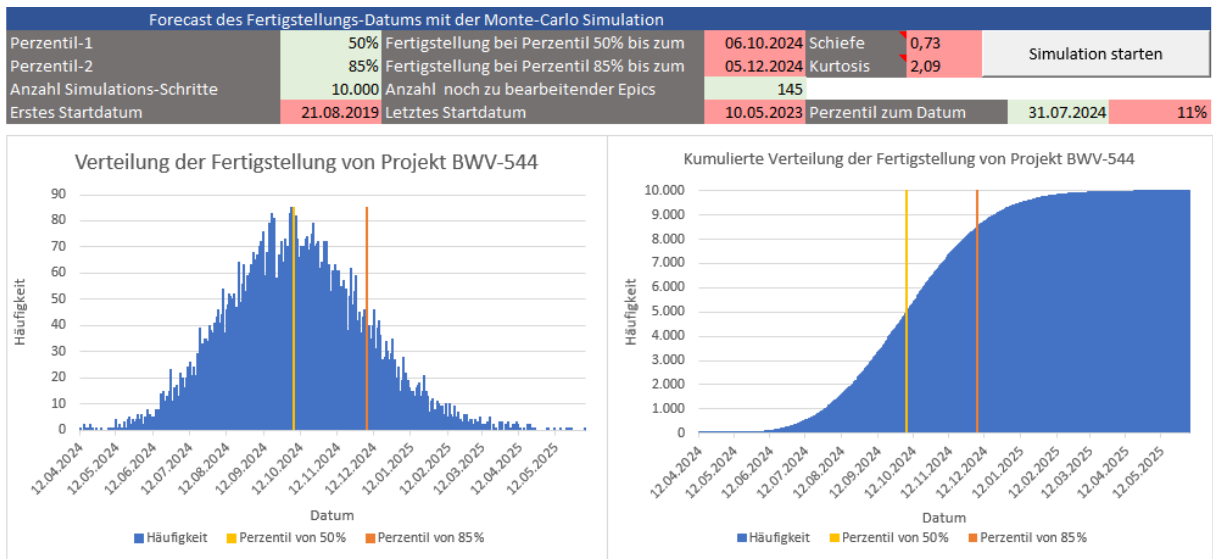


Abbildung 7: simulierter Forecast für das Großprojekt

» Bei diesem nicht vorteilhaften Prozess ist bei einem stabilen Backlog mit 145 Epics mit einer Fertigstellung bei einer Wahrscheinlichkeit zwischen 50% und 85% von 06.10.2024 bis 05.12.2024 zu rechnen. Bei Analyse der Entwicklung des Durchsatzes über die Entwicklungszeit kann man erkennen, dass der Durchsatz Ende 2021 geringer als danach war:

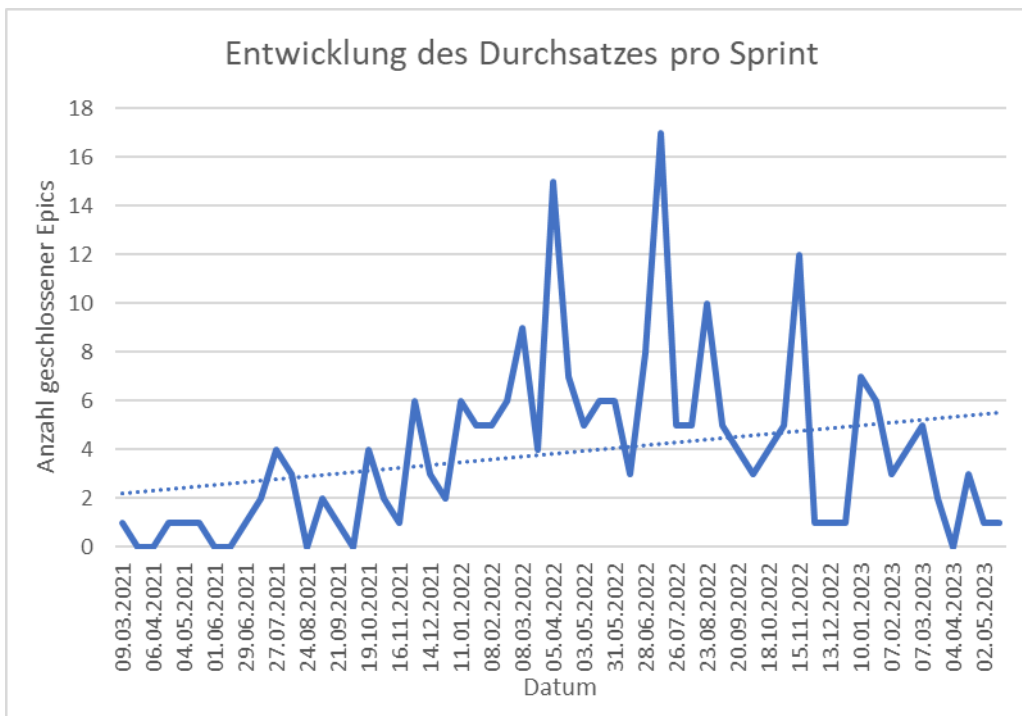


Abbildung 8: Entwicklung des Durchsatzes über die Zeit

Für einen realistischeren Forecast ist es daher sinnvoller, ein eingeschränktes Input-Datenset zu verwenden.

Zur Wahl eines Zeitraumes bestehender Daten für eine aktuellere Forecast-Simulation sollen zuerst zwei Aspekte beleuchtet werden:

Der Durchsatz zeigt seit Anfang 2022 einen Abwärtstrend:

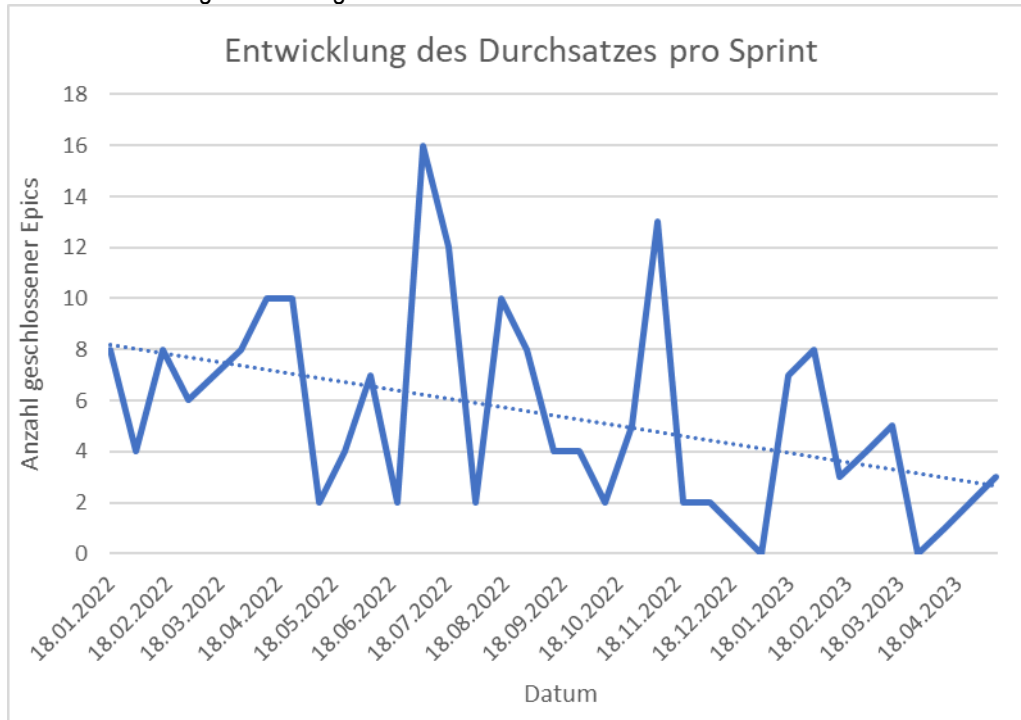


Abbildung 9: Entwicklung des Durchsatzes seit Anfang 2022

Weiterhin kann man an dem sog. Aging-Chart erkennen, dass das Projekt aktuell ein Problem mit der Fluss-Effizienz hat:

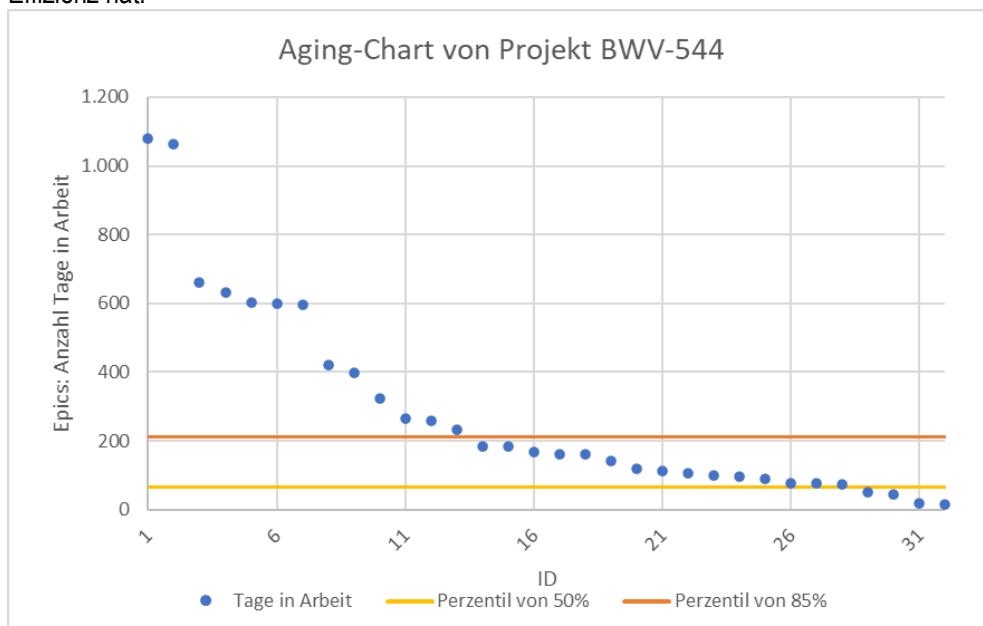


Abbildung 10: Aging-Chart: aktuelle Dauer der Epics in Arbeit

Die Cyclictimes bei den Perzentilen von 50% und 85% beziehen sich auf die Perzentile vom 01.01.2022 bis 02.05.2023. Es ist klar zu erkennen, dass nur 4 der 32 Epics unter dem Perzentil von 50% liegen, also sich aktuell noch unter 64 Tagen in Arbeit befinden. 13 von 32 Epics liegen sogar über dem Perzentil von 85% (212 Tage), sind also schon länger als 212 Tage in Bearbeitung. Dies spiegelt genau auch die Situation wider, dass das Projekt sehr stark mit Planungen und Problemen durch Einschränkungen der Architektur belastet ist und man somit auch die Fluss-Effizienz gänzlich außer Acht gelassen hat.

Eine erneute Forecast-Simulation soll mit den Daten der ab Anfang September 2022 abgeschlossenen Epics durchgeführt werden. Hierzu kann man in dem Tool auch einen Zeitraum definieren.

Mit diesem Datensatz ergeben sich folgende Auswertungen:

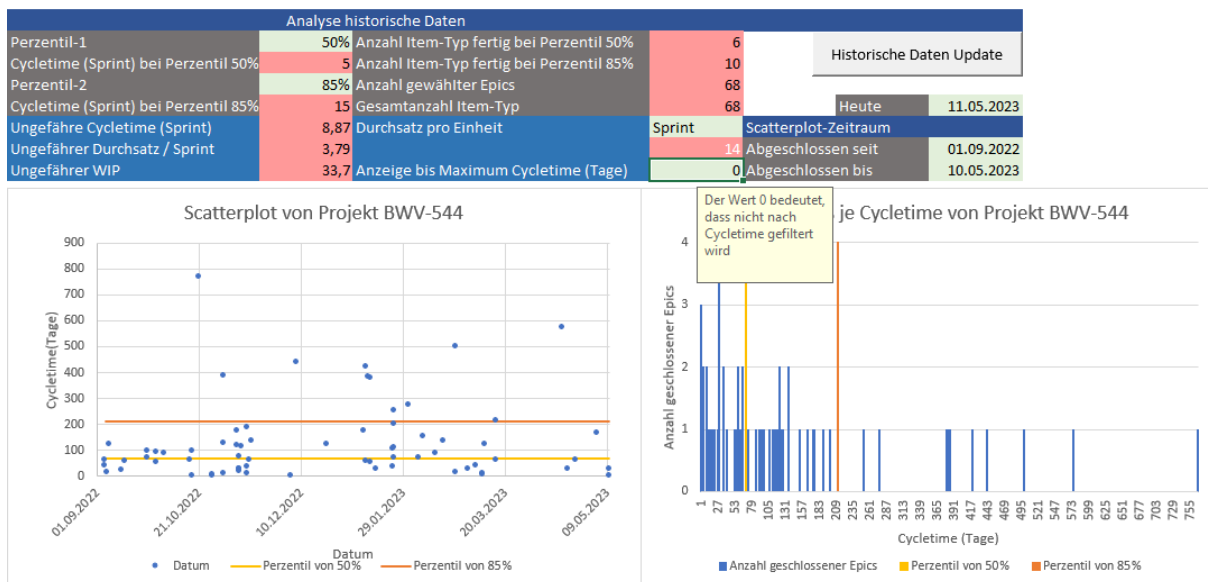


Abbildung 11: Auswertung aller Epics mit Enddatum >= 01.10.2021

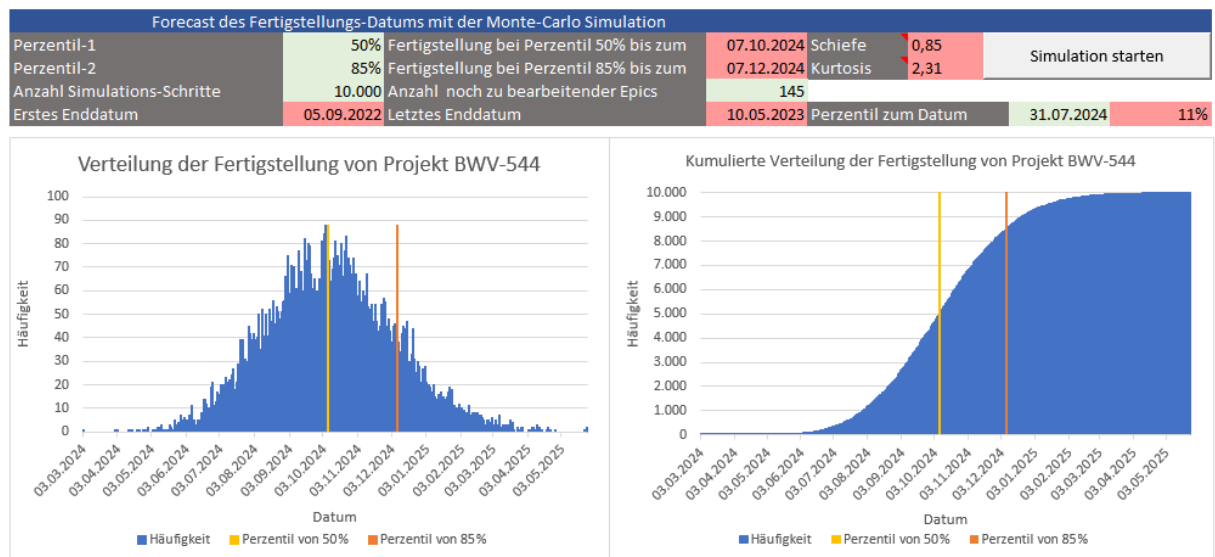


Abbildung 12: simulierter Forecast mit Epics, die ab September 2021 fertiggestellt wurden

Die Schiefe ist ein Kriterium, um zu prüfen, ob eine symmetrische Verteilung vorliegt, das heißt die Wahrscheinlichkeit für positive und negative Abweichungen gleich hoch sind. Da die Normalverteilung eine symmetrische Verteilung ist, hat sie eine Schiefe von null. Eine linksschiefe Verteilung hat einen negativen Wert, eine rechtsschiefe Verteilung einen positiven Wert. Die Schiefe von 0,79 bedeutet, dass spätere Fertigstellungstermine etwas wahrscheinlicher sind als Frühere.

Die Kurtosis beschreibt die Wölbung der Verteilung und hat für die Normalverteilung den Wert 3. Die Wölbung bringt zum Ausdruck, wie spitz eine Verteilung zuläuft. Eine Wölbung > 3 bedeutet, dass Extremwerte in den äußerst linken und rechten Enden der Verteilung häufiger vorkommen, als es die Normalverteilung annimmt, was als „fat tails“ bezeichnet wird. Bei einer Kurtosis von 2,17 ist mit weniger Ausreißern zu rechnen.

Gegenüberstellung der Messergebnisse:

Messwert	Alle Epics	Nur Epics ab September 2022
Cycletime mit Perzentil 50% (Tage)	<= 68	<= 70
Cycletime mit Perzentil 85% (Tage)	<= 240	<= 211
Durchschnittliche Cycletime (in Wochen)	16,39	17,75
Durchschnittlicher Durchsatz / Woche	1,96 Epics	1,9 Epics
Durchschnittlicher WIP	32,2 Epics	33,7 Epics
Fertigstellung bei 50% Perzentil	08.10.2024	07.10.2024
Fertigstellung bei 85% Perzentil	09.12.2024	07.12.2024

Tabelle 1: Gegenüberstellung der Messwert mit gefilterten Epics mit Epic-Fertigstellung ab 01.09.2022

Mit der Simulation kann weiterhin die Anzahl möglicher Fertigstellungen von Work-Items (hier Epics) berechnet werden:

Forecast der Anzahl fertiggestellter Epics bis zum gewünschten Tag mit der Monte-Carlo Simulation			
Bis Datum	31.07.2024	Entspricht 448 Tagen	
Maximal fertiggestellte Epics bei Perzentil 50%	125	Perzentil bei Anzahl Epics	145
Maximal fertiggestellte Epics bei Perzentil 85%	110		9%
		Schiefe	0,84
		Kurtosis	2,17

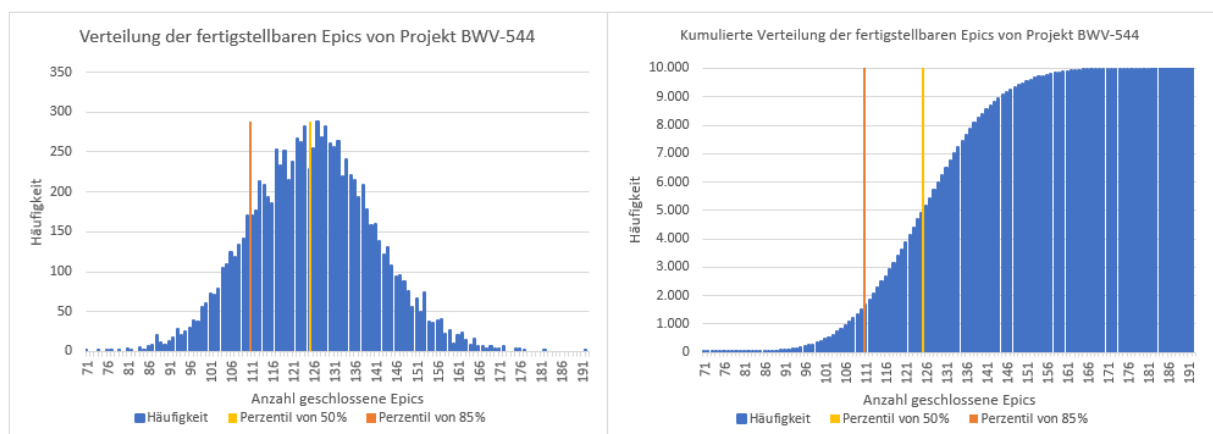


Abbildung 13: Simulation einer möglichen Anzahl Fertigstellungen von Epics

Wenn man die Ausreißer mit einer Cycletime >= 300 Tage herausfiltert – hier handelt es sich um 22 Epics –, kann auch die Verteilung der Cycletime genauer analysiert werden.



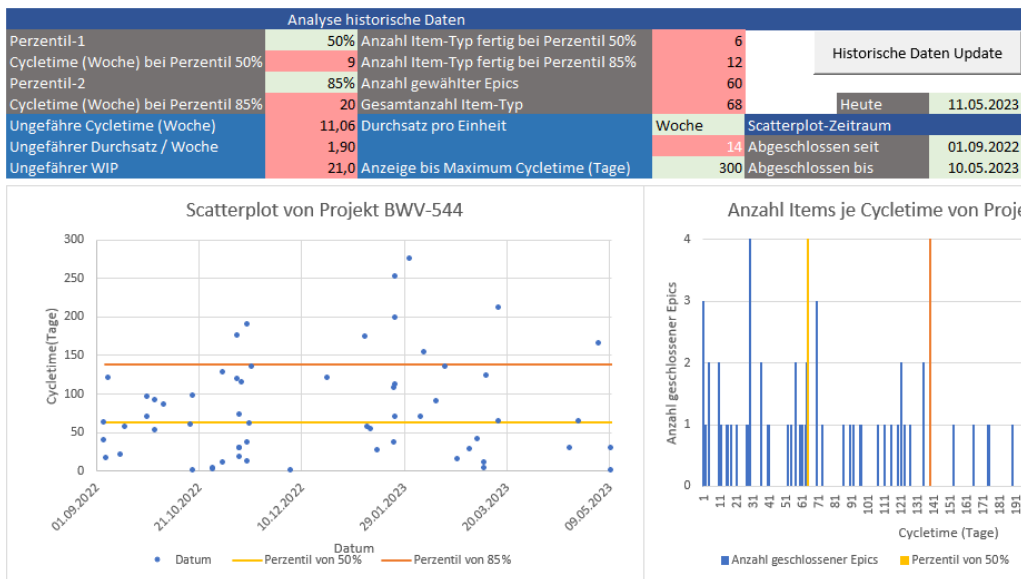


Abbildung 14: Auswertungen, nachdem Epics mit einer Cycletime  $\geq 300$  herausgefiltert wurden

Messwert	Alle Epics ab September 2021	Epics mit Cycletime < 300 Tage
Cycletime mit Perzentil 50% (Tage)	$\leq 70$	$\leq 64$
Cycletime mit Perzentil 85% (Tage)	$\leq 211$	$\leq 138$
Durchschnittliche Cycletime (in Wochen)	17,75	11,06
Durchschnittlicher Durchsatz / Woche	1,9 Epics	1,9 Epics
Durchschnittlicher WIP	33,7 Epics	21,0 Epics
Fertigstellung bei 50% Perzentil	07.10.2024	30.11.2024
Fertigstellung bei 85% Perzentil	07.12.2024	04.02.2025

Tabelle 2: Gegenüberstellung der Messwert mit gefilterten Epics anhand Cycletime

Folgende Auffälligkeiten sind zu beobachten:

- Durch eine Prozessverbesserung sollte es gelingen, Ausreißer zu entfernen. Dadurch verringert sich der WIP um ca. ein Drittel. Gleichzeitig reduziert sich die durchschnittliche Cycletime um ein Viertel.
- Im Zeitraum zwischen Januar 2022 und November scheinen mehr Epics pro Woche fertiggestellt worden zu sein als davor bzw. danach (dies ist auch im Durchsatz-Diagramm zu erkennen). In der Tat wurde während des Zeitraums März 2022 bis Oktober 2022 auch ein Augenmerk auf die Fokussierung bzgl. Cycletime und WIP gelegt. Diese Fokussierung wurde aufgrund von Neuplanungen und einhergehenden Schätzungen mit viel Aufwand aufgegeben. Hätte man auf Schätzungen verzichtet, dann hätte man auch besser den Fokus halten können.
- Die Verteilung der Häufigkeiten der Cycletimes:
  - » Die Verteilung zeigt eindeutig keine Normalverteilung. Diskrete Forecasts basieren oftmals auf Mittelwerten  $\pm$  Standardabweichung, sprich Normalverteilungen. Diese Methode des diskreten Forecasts würde daher ein völlig verzerrtes Ergebnis ergeben.
  - » Die Verteilung entspricht nicht ganz einer typischen Cycletime-Verteilung bei Wissensarbeiten, denn es gibt hier noch zu viele Cycletimes im Long-Tail Bereich.



- Prognostizierte Fertigstellungs-Bandbreite: die beiden Szenarien haben fast identische prognostizierte Fertigstellungs-Bandbreiten, was auf den ersten Blick verwirrt. Diese Ähnlichkeit liegt daran, dass der probabilistische Forecast auf dem Durchsatz basiert, und dieser unterscheidet sich bei beiden Szenarien nur um 0,11 Epics pro Woche. Der Durchsatz beim 1. Szenario ist höher, und damit ist mit einem früheren Fertigstellungstermin zu rechnen. Im zweiten Szenario wurden 10 Epics, deren Cycle-Time größer als 300 Tage waren, aus der Input-Liste entfernt, obwohl diese 10 Epics in Realität Aufwand gekostet haben. Wären diese 10 Langläufer in Realität frühzeitig aus dem Prozess entfernt worden, oder hätten diese schneller abgearbeitet werden können, dann hätte sich auch der Durchsatz erhöht.

Um die Auffälligkeiten der langen Cycle-times frühzeitig zu erkennen, sollten auf jeden Fall CFDs (Cumulative Flow Diagram) eingesetzt werden.

### 5. Visualisierung der Prozess-Stabilität mit dem „Cumulative Flow Diagram (CFD)“

Bei diesen Diagrammen handelt es sich ebenfalls wie bei Little’s Law um eingehende und ausgehende Items, nur eben in Echtzeit. CFDs visualisieren auf eine exzellente Art den Fluss der Arbeit bzw. der Items durch den Prozess. Mit dieser Visualisierung kann man (und muss man) sehr früh auf Wartezeiten und Blockaden aufmerksam werden, um diese zu beheben und die Flow-Efficiency zu steigern.

Cumulative Flow Diagram (CFD)			
Gesamter Zeitraum		Gewählter Zeitraum	
Start	25.05.2020	Start	25.05.2020
Ende	09.05.2023	Ende	09.05.2023
Durchschnittlicher WIP (Monat)	29,65	Durchschnittlicher WIP (Monat)	29,65
Durchschnittliche Cycletime (Monat)	4,80	Durchschnittliche Cycletime (Monat)	4,80
Durchschnittlicher Durchsatz / Monat	6,17	Durchschnittlicher Durchsatz / Monat	6,17
Durchschnittliche Ankunftsrate / Monat	7,06	Durchschnittliche Ankunftsrate / Monat	7,06
Eingegangene Epics Gesamt	255	Eingegangene Epics Gesamt	255
Ausgegangene Epics Gesamt	222	Ausgegangene Epics Gesamt	222

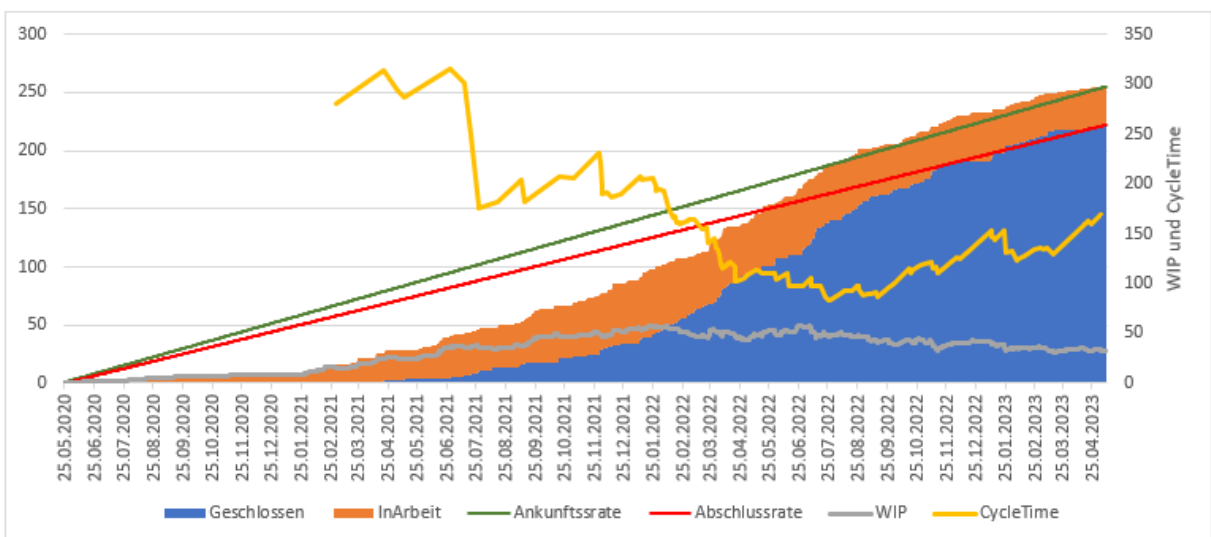


Abbildung 15: CFD des Großprojektes ab Umstieg auf Jira

Mit einem CFD können, wie in Abbildung 15 zu sehen, mehrere Prozess-Merkmale erkannt werden:

- Im Zeitraum von Mai 2020 bis Mai 2021 wurden so gut wie keine Epics geschlossen (blaue Fläche).
- gleichzeitig wurden Epics in Arbeit genommen (orange Fläche)
- so dass sich in dieser Zeit der WIP (graue Linie) bis dahin auf 35 in Arbeit befindlichen Epics vergrößerte, was natürlich in Summe für 4 Teams viel zu hoch ist; ein fehlender Fokus ist festzustellen, es gab zu viele Abhängigkeiten zwischen den Teams, was zu Wartezeiten führte, auch Blockaden auf Grund fehlender Anforderungsabstimmungen traten auf. Der WIP wird berechnet als vertikale Distanz zwischen den oberen Linien der in Arbeit befindlichen und der abgeschlossenen Epics zu einem jeweiligen Datum (Wert auf der X-Achse).
- Die durchschnittliche Cycletime (gelbe Linie) im Prozess hat kontinuierlich bis Juni 2022 abgenommen, was dadurch begründet ist, dass ab Januar 2022 mehr Epics geschlossen werden konnten als mit neuen Epics begonnen wurden. Die Prozess-Cycletime wird berechnet als horizontale Distanz zwischen den oberen Linien der in Arbeit befindlichen und der abgeschlossenen Epics bei einem jeweiligen Y-Wert (Anzahl Eics).
- Im gesamten Zeitverlauf liegt jedoch die Abschlussrate (rote Gerade) unter der Ankunftsrate (grüne Gerade). Die Ankunftsrate ist die durchschnittliche Anzahl angefangener Epics pro Zeiteinheit, wogegen die Abschlussrate die durchschnittliche Anzahl abgeschlossener Epics je Zeiteinheit ist.
- Diesen Sachverhalt sieht man auch als Zusammenfassung über dem CFD:
  - » Durchschnittlicher Durchsatz (abgeschlossene Epics) pro Monat: 6,17 Epics
  - » Durchschnittliche Ankunftsrate (neu begonnene Epics) pro Monat: 7,06 Epics
  - » Eingegangene Epics gesamt (also in Arbeit genommene Epics): 255 Epics
  - » Ausgegangene Epics gesamt (abgeschlossene Epics): 222 Epics
  - » Es wurden also mehr Epics in das System gegeben, als abgearbeitet werden konnten, was immer ein Indiz für hohe Wartezeiten in den Epics ist und somit den Prozess verlangsamt

Man kann sich natürlich auch in ein CFD herein zoomen; als Start-Termin wird der 01.06.2021 gewählt:

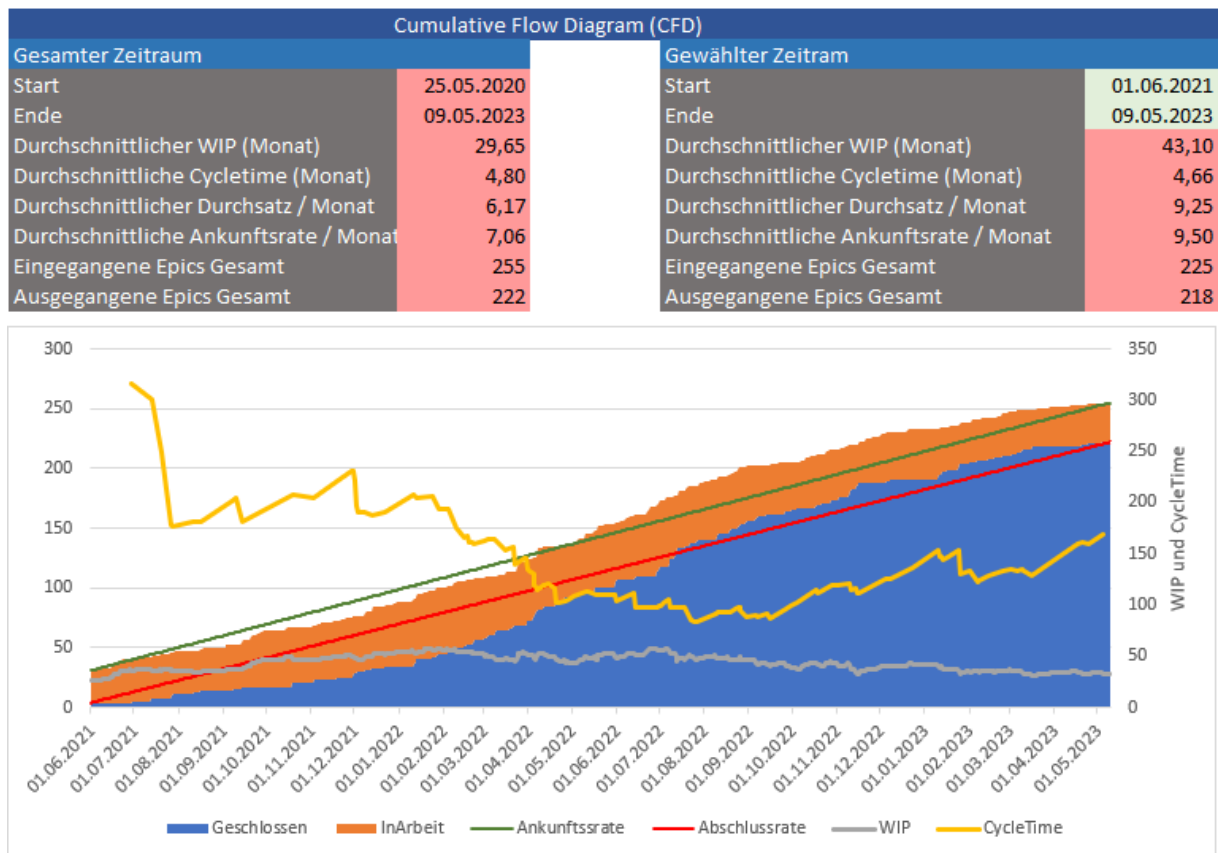


Abbildung 16: Zoom in das CFD

Im Zeitraum vom 01.06.2021 bis 10.05.2023 ist die Ankunftsrate ungefähr gleich zu der Rate der abgeschlossenen Epics. Von einer Prozess-Verbesserung kann hier trotzdem nicht gesprochen werden, da sich die durchschnittliche Cycletime ab September 2022 wieder vergrößert hat.

## 6. Schätzen von User-Stories ist nicht zielführend

In Sprint-Plannings und weiteren Planungsrunden wird viel Zeit und Aufwand für das Schätzen von User-Stories und Aufgaben aufgewendet. Die allgemeine Auffassung bzgl. des Schätzens lautet oftmals wie hier aus IAPM entnommen:

*Eine wichtige Grundlage für jede weitere Planung ist das Abschätzen des Arbeitsaufwandes, der sich hinter den einzelnen Einträgen im Product Backlog verbirgt. Damit lässt sich ein erster Eindruck vermitteln, welchen Umfang die Arbeiten im Projektverlauf haben könnten. Dabei wird darauf verzichtet, konkrete Vorhersagen zu machen, wie hoch beispielsweise der Aufwand in Arbeitsstunden sein könnte. Vielmehr geht es darum, relative Größen zu ermitteln, das heißt, die Komplexität der einzelnen Einträge wird im Verhältnis zueinander geschätzt. Durch die relativen Aufwandsgrößen ist es einfacher, die Release-Pläne an geänderte Randbedingungen, die die Velocity (Entwicklungsgeschwindigkeit) verändern, anzupassen, und die Auswirkungen der Veränderungen zu bewerten.*

Der Fokus wird, statt auf genaue Zahlen auf eine gewollte Ungenauigkeit und damit auf nicht lineare Zahlenskalen gesetzt. Zusätzlich wird die Abschätzung "aus dem Bauch" heraus vorgenommen.

Wenn jedoch das Management fragt, wann ein Release oder das Projekt beendet sein wird, so kann man schwerlich mit Story-Points argumentieren. Weitaus gravierender jedoch ist die tatsächliche Dauer, die Stories mit unterschiedlichen Story-Points einnehmen. Hier gibt es so gut wie nie eine Korrelation zwischen Dauer und Story-Points.

Dies soll anhand einer Auswertung von 4 Teams des Großprojektes veranschaulicht werden:

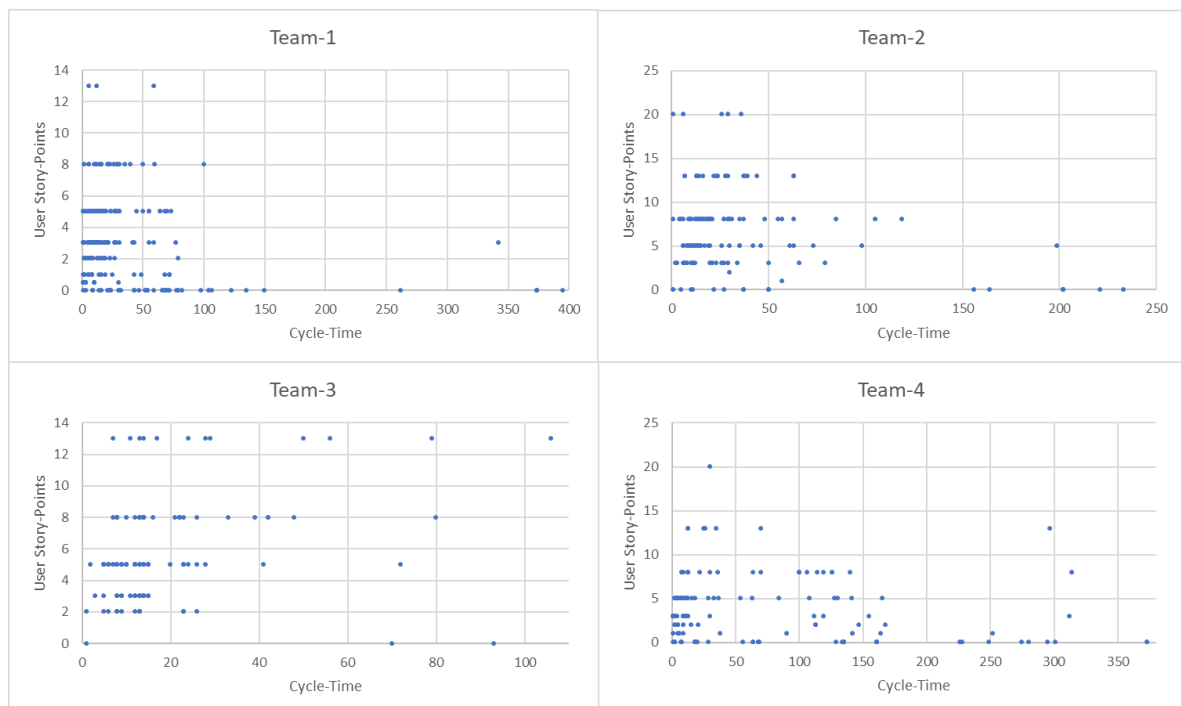


Abbildung 17: Dauer von User-Stories mit unterschiedlichen Story-Points

Drei der Teams nutzten Stories für Maßnahmen, um z.B. technische Schulden abzubauen. Solche Stories wurden mit 0 Story-Points bewertet. Aber auch diese benötigen Zeit und Aufwand, der bei Velocity-Auswertungen vernachlässigt wird.

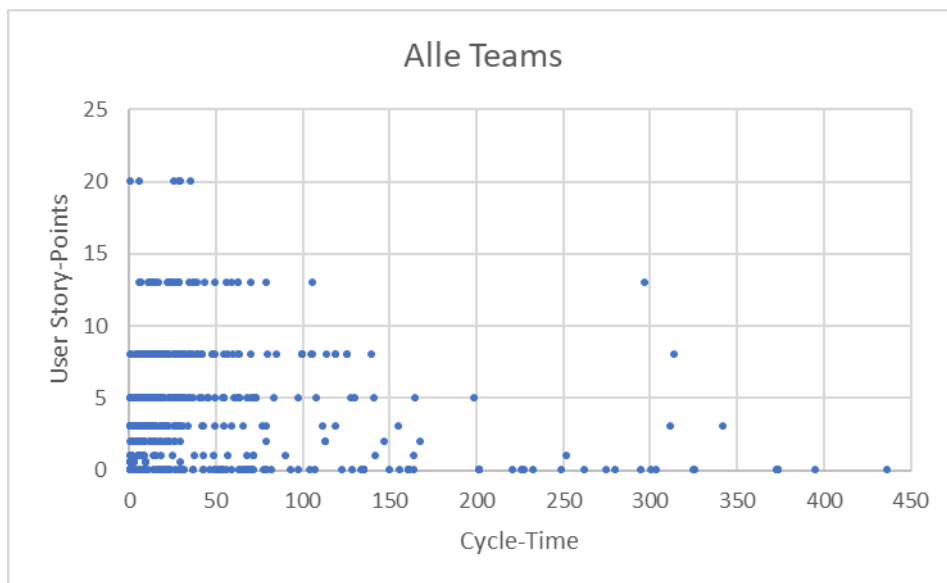


Abbildung 18: Zusammenfassung der Auswertung Story-Point vs. Dauer von User-Stories über alle 4 Teams

Es ist klar erkenntlich, dass Stories mit z.B. 5 Story-Points teilweise länger dauern als Stories mit 13 Story-Points.

Eine numerische Auswertung verdeutlicht dies:

Story-Points	Mittlere Cycle-Time (Tage)					Anzahl User-Stories					
	Alle Teams	Team-1	Team-2	Team-3	Team-4	Alle Teams	Team-1	Team-2	Team-3	Team-4	
0	140	111	99	336	157	121	49	21	8	43	
1	39	21	57		79	31	21	1	0	9	
2	22	12	30	12	60	50	26	1	13	10	
3	24	21	24	11	42	113	58	18	17	20	
5	24	22	27	15	36	162	54	39	37	32	
8	35	26	27	22	77	112	22	43	27	20	
13	37	26	28	36	68	40	3	18	12	7	
20	21		20		30	6	0	5	0	1	
Durchschnitt	48	37	37	40	85	Gesamt	635	233	146	114	142

Abbildung 19: Numerische Auswertung Story-Points versus Dauer von User-Stories

Im Durchschnitt benötigen Stories mit 1 Story-Point mehr Tage bis zur Fertigstellung als Stories mit 13 oder 20 Story-Points.

## 7. Epilog

Für eine Prognose der Fertigstellung von Releases oder eines Projektes können Schätzungen mit beispielsweise Story-Points zu erheblichen Fehleinschätzungen führen. Eine lineare Prognose – auch mit Standardabweichungen - anhand von Velocity und ausstehender Gesamtanzahl von Schätzwerten wie Story-Points führt oftmals zu falschen Ergebnissen. Für ein realistisches Forecasting und zur Steigerung der Arbeitsfluss-Effizienz sind folgende Punkte wesentlich:

- Nutze die Auswertung von Einzelitems in Form von Scatter-Plot und Histogramm, um die Verteilung und die Perzentile der Dauer der Work-Items (z.B. Epics oder User-Stories) zu analysieren. Es fällt auf, dass die Dauer der Abarbeitung der Work-Items nicht normalverteilt ist und daher mit Mittelwerten nicht

gerechnet werden kann. Viel wichtiger ist es, die Dauer anhand von Perzentilen also Wahrscheinlichkeiten zu berechnen. Diese Analyse sollte mindestens einmal die Woche durchgeführt werden.

- Nutze das Aging-Chart, um täglich die Dauer in Arbeit befindlicher Work-Items zu beobachten und zu analysieren, warum bestimmte Work-Items schon lange in Arbeit sind, um entsprechend schnell Maßnahmen zu ergreifen. Nur so entsteht Fokus auf die aktuelle Arbeit.
- Nutze das CFD, um die Ankunfts- und Abflugrate (also Anzahl gestarteter und beendeter Work-Items in einem Zeitraum zu analysieren, aber auch den WIP, den Durchsatz und die Cycle-Time.
- Nutze all diese Informationen, um den Prozess zu stabilisieren und zu optimieren. Nur mit einem stabilen Prozess können auch Prognosen erstellt werden. Beobachte den WIP und reduziere den WIP, um somit die Cycle-Time zu mindern und den Durchsatz zu erhöhen. Nutze auch diese Informationen, um Prozess-Regeln zu definieren bzw. zu schärfen.
- Nutze die Monte-Carlo-Simulation, um eine probabilistische Wahrscheinlichkeitsverteilung von Fertigstellungsterminen oder möglicher abzuschließender Anzahl Work-Items zu erhalten.
  - » Wähle hierzu den sinnvollsten Zeitrahmen historischer Daten aus.
  - » Definiere die Perzentile, die einen Orientierungsrahmen geben.
  - » Mache die ersten Forecasts und experimentiere mit den Input-Daten
  - » Es genügen auch wenige Input-Daten, somit kann bereits frühzeitig mit einer MCS begonnen werden. Denn es gilt die „Rule of Five“ von Douglas Hubbard: *„There is a 93,75% chance that the median of a population is between the smallest and largest values in any random sample of five from that population“*.
- Führe die MCS immer wieder durch, sobald neue Informationen (Abschluss von Work-Items) vorliegen. Somit können frühzeitig Maßnahmen ergriffen werden, wie z.B.:
  - » Nimm Scope aus dem Backlog für das Release oder Projekt
  - » Ändere den Fertigstellungs-Termin
  - » Füge weitere Mitarbeiter hinzu. Wenn man nicht früh genug die Teams verstärkt, wird Brooks Law erbarmungslos zuschlagen: *„Adding people to a late project makes it even later“*
  - » Akzeptiere das erhöhte Risiko, den Termin nicht einhalten zu können
  - » Oder eine Kombination von den vier Maßnahmen

Es sei nochmals darauf hingewiesen, dass solche probabilistischen Forecasts permanent neu ausgewertet werden müssen, denn im Laufe der Zeit ändern sich zum einen die Statistiken der in Summe abgeschlossenen Work-Items und zum anderen kann sich der Backlog, also die noch ausstehenden bzw. in Arbeit befindlichen Items, in seiner Anzahl Items ändern. Für Forecasts sollte eine Monte-Carlo-Simulation herangezogen werden, und der Forecast sollte immer basierend auf Perzentilen kommuniziert werden, denn die Perzentil-Linien und die assoziierten Bereiche und Wahrscheinlichkeiten repräsentieren den Forecast für den Backlog. Es sollte nie eine deterministische, lineare Prognose angegeben werden, da dies zu falschen Erwartungen führen wird.

Beim Forecasting geht es nicht nur darum, Daten zu sammeln, um Prognosen zu erstellen – es geht um die Maßnahmen, die man jeden Tag ergreifen soll, um die Vorhersehbarkeit des Prozesses sicherzustellen.

Das Wichtigste, was man tun kann, um die Prognosen für einzelne Items zu verbessern, besteht darin, proaktiv zu bestimmen, wie lange es dauert, bis die Items fertig sind.

## 8. Abkürzungen

Abkürzung	Beschreibung
<b>CFD</b>	Cumulative Flow Diagram
<b>Cycle Time</b>	Zykluszeit: Ist die Zeit, die vom Beginn bis zur Fertigstellung einer Aufgabe benötigt wird. Die Cycle Time ist von der Lead Time zu unterscheiden.
<b>Durchsatz</b>	Durchschnittliche Anzahl abgeschlossener Items je Zeiteinheit
<b>Forecast</b>	Vorhersage, Prognose
<b>Kurtosis</b>	Die Kurtosis beschreibt die Wölbung der Verteilung und hat für die Normalverteilung den Wert 3. Die Wölbung bringt zum Ausdruck, wie spitz eine Verteilung zuläuft. Eine Wölbung > 3 bedeutet, dass Extremwerte in den äußerst linken und rechten Enden der Verteilung häufiger vorkommen, als es die Normalverteilung annimmt, was als „fat tails“ bezeichnet wird. Bei einer Kurtosis von 2,17 ist mit weniger Ausreißern zu rechnen.
<b>Lead Time</b>	Durchlaufzeit: Ist die Zeit, die benötigt wird, um eine Aufgabe seit ihrem Bekanntwerden fertig zu stellen. Dies kann beispielsweise der Zeitraum sein, in welchem ein Kunde seinen Auftrag kommuniziert, bis zu dem Auslieferungszeitpunkt des fertiggestellten Produktes.
<b>MCS</b>	Monte-Carlo-Simulation
<b>OM</b>	Operations Management
<b>Schiefe</b>	Die Schiefe ist ein Kriterium, um zu prüfen, ob eine symmetrische Verteilung vorliegt, das heißt die Wahrscheinlichkeit für positive und negative Abweichungen gleich hoch sind. Da die Normalverteilung eine symmetrische Verteilung ist, hat sie eine Schiefe von null. Eine linksschiefe Verteilung hat einen negativen Wert, eine rechtsschiefe Verteilung einen positiven Wert. Die Schiefe von 0,79 bedeutet, dass spätere Fertigstellungstermine etwas wahrscheinlicher sind als Frühere.
<b>WIP</b>	Work in Progress oder besser gesagt „Work in Process“: „Umlaufbestand“ oder „Ware in Arbeit“. Darunter versteht man Materialbestände bzw. laufende Arbeitsaufträge, die den Fertigungsaufträgen und nicht dem Lagerbestand zugeordnet sind.

Tabelle 3: Abkürzungen