Homepageveröffentlichung unbefristet genehmigt für Zedas GmbH /
Rechte für einzelne Downloads und Ausdrucke für Besucher der Seiten
genehmigt von DVV Media Group, 2018.

Datengestützte Entscheidungen zum Lokeinsatz

Einsparungspotenziale mithilfe von Big Data erkennen und nutzen – ein praktisches Beispiel

ULRICH LIESKE

Die effiziente Nutzung der Produktionsmittel ist ein wesentlicher Kostenfaktor für jedes Unternehmen. Das gilt auch für die Lokomotiven der Eisenbahnverkehrsunternehmen (EVU). Der Grat für operative und strategische Entscheidungen zu deren Einsatz ist schmal und wird zumeist von den Erfahrungen und dem Bauchgefühl einzelner Personen bestimmt. Neue Technologien, wie Big Data, Advanced Analytics und Operative Intelligenz, bieten neue Ansätze, Entscheidungen auf der Basis von Daten und deren Auswertung zu treffen. Anhand eines konkreten Praxisbeispiels wird beschrieben, wie sich mithilfe dieser Technologien erhebliche Kostensenkungspotenziale ermitteln und umsetzen lassen.

Die Aufgabe verstehen

Ausgangspunkt war die in einem EVU allgemein gestellte Frage nach der zeitlichen Auslastung der eigenen Lokomotiven. Im Gespräch mit den Beteiligten kristallisierten sich schnell zwei grundlegende Sichtweisen auf das Thema heraus. Die Kollegen aus der Lokeinsatzplanung bezogen die Fragestellung auf die zeitliche Verfügbarkeit der Loks. Die Lok fährt – sie ist „Im Einsatz". Die Lok ist abgestellt, kann aber eingesetzt werden – sie ist „Verfügbar" oder die Lok ist „Nicht Verfügbar", weil sie eine Störung hat oder sich in der Werkstatt befindet. Andere sahen in der Frage nach der zeitlichen Auslastung vorrangig deren produktive Verwendung. Fährt eine Lok per Auftrag „Geld ein", dann ist sie „Produktiv". Bei einer Überführung zum Startpunkt des nächsten Auftrags fährt sie „Nicht Produktiv". Eine Lok ohne Auftrag, die nicht fährt, ist „Abgestellt". Einer Zuführung in die bzw. einer Rückführung aus der Werkstatt wurde der Status „Rangiert" zugeordnet. Die Zeit der Entstörung oder Reparatur fiel unter das Schlagwort „Werkstatt".

Um beiden Seiten gerecht zu werden, wurden zunächst die sinnvollen Kombinationen von Einsatz- und Produktionsstatus ermittelt (Abb. 1).

Jeder Kombination ließen sich danach Zeiten zuordnen, deren Ermittlung Ergebnis der Datenaufbereitung sein sollten. Auch der Fall, dass aus den Daten keine Zuordnung möglich ist, wurde in der Zeitendefinition berücksichtigt („Sonstige Zeit"). Jede Zeitendefinition erhielt zudem ein handhabbares Kürzel.

Im Ergebnis wurden zehn verschiedene Zeiten vereinbart, deren Definition und Bedeutung aus betrieblicher Sicht als Vorgabe für die Datenanalyse dienten (Tab. 1).

Nach Wunsch des Auftraggebers sollten die Anteile dieser Zeiten für jede Lok tagesaktuell zur Verfügung stehen. Für die Auswahl des Analysezeitraums waren nicht nur Stunde, Kalendertag, -monat und -jahr vorgegeben, sondern auch Schicht, Produktionstag, -monat und -jahr.

Notwendige Daten

Voraussetzung für jede Analyse von Lokeinsatzzeiten sind operative Prozessdaten mit Zeitstempel, Objektidentifikation (Loknummer), Ortsangabe (Gleis, Bahnhof) und Zusatzinformationen (Transportgut, Werkstatt). Eine hohe Qualität der Analyse ist dabei nur mit einer zuverlässigen, technischen Erfassung und Speicherung dieser Daten möglich. Im konkreten Fall stehen elektronische Gleisfortschaltungsmeldungen zur Verfügung, die zeitnah übermittelt werden. Sie enthalten u. a. die Loknummer, das Start- und Zielgleis, einen Transportcode und den Zeitstempel der Einfahrt in das Zielgleis. Die Einfahrtzeit in das Zielgleis ist gleich der Ausfahrtzeit aus dem Startgleis. Die Differenz der Zeitstempel zweier aufeinanderfolgenden Gleisfortschaltungen einer Lok ergibt somit die Aufenthaltszeit im Startgleis der zweiten Meldung.

Aus dem Instandhaltungssystem zedas asset und dem Auftragssystem zedas cargo des Kunden konnten notwendige Zusatzinformationen (Transportgut, Fahrzeugstammdaten, Werkstattinformationen) ebenfalls elektronisch abgegriffen werden. Verschiedene Stammdatentabellen ergänzen die operativen Daten. Um Be- und Entladezeiten ermitteln zu können, wurden die Be- und Entladegleise in der Stammdatentabelle „Ereignisgleise" erfasst. Die Definition der Schichtwechselzeiten und damit der Produktionstage abweichend vom Kalendertag befinden sich in der Stammdatentabelle „Schicht-Definition".

Lösungsansätze zur Zeitenermittlung

Die aus den operativen Gleisfortschaltungen gewonnene Aufenthaltszeit im Gleis gibt zunächst keine Auskunft darüber, welchen Anteil die Fahrtzeit, die Standzeit und eventuell die Ladezeit im Gleis haben. Wie lassen sich diese Anteile ermitteln, ohne dass weitere Informationen zur Verfügung stehen?

Fahrtzeit vs. Standzeit

Den Schlüssel zur Antwort liefert eine statistische Auswertung der historischen Gleisfortschaltungsdaten. Aus ihnen werden alle eingetretenen Kombinationen von Gleisbefahrungen, bestehend aus Richtungsgleis, Startgleis, Zielgleis sowie Transportgut im Start- und Zielgleis, ein Jahr rückwirkend ermittelt. Das Richtungsgleis ist dabei das vor dem Startgleis befahrene Gleis. Der Wechsel des Transportguts zwischen Start- und Zielgleis impliziert eine Be- bzw. Entladung im Startgleis. Zu jeder stattgefundenen Fahrt von

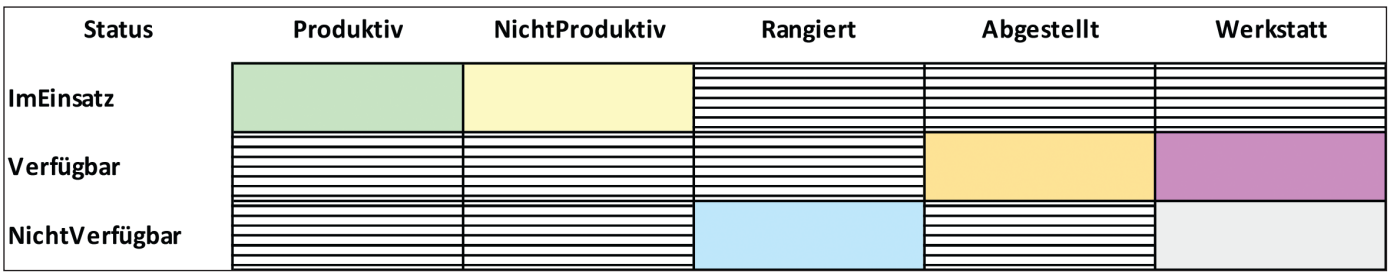

Abb. 1: Relevante Kombinationen aus Produktiv- und Einsatzstatus von Lokomotiven

Homepageveröffentlichung unbefristet genehmigt für Zedas GmbH /
Rechte für einzelne Downloads und Ausdrucke für Besucher der Seiten
genehmigt von DVV Media Group, 2018.

Einsatzstatus	Produktionsstatus	Zeiten	Zeitkürzel	Zeitaggregierung
Im Einsatz	Produktiv	Fahrzeit Produktiv	FZP	Produktive Einsatzzeit
Im Einsatz	Produktiv	Standzeit Produktiv	SZP	
Im Einsatz	Produktiv	Be-/Entladezeit Produktiv	BEZ	
Im Einsatz	Nicht Produktiv	Fahrzeit Nicht Produktiv	FZNP	Nicht Produktive Einsatzzeit
Im Einsatz	Nicht Produktiv	Standzeit Nicht Produktiv	SZNP	
Verfügbar	Abgestellt	Fahrbetriebliche Abstellzeit	AZFB	Abstellzeit
Verfügbar	Werkstatt	Abstellzeit nach Werkstatt	AZWS	
Nicht Verfügbar	Rangiert	Rangierzeit (Zuführung Werkstatt)	RZ	Werkstattzeit
Nicht Verfügbar	Werkstatt	Werkstattzeit (Geplant und Störung)	WSZ	
Status Nicht Zuordenbar		Ohne Status Zeit	OSZ	Ohne Status Zeit

Tab. 1: Vorgaben zur Zeitenermittlung

Richtungsgleis über Startgleis in das Zielgleis mit jeweiligem Transportgut wird die Gleisaufenthaltszeit im Startgleis ermittelt. Der statistische Zentralwert (Median) dieser Zeiten (Abb. 3) ergibt die „Typische Gleisaufenthaltszeit" im Startgleis jeder dieser Kombinationen. Diese Informationen werden monatlich neu berechnet und in der dynamisch generierten Stammdatentabelle „Typische Gleiszeiten" abgespeichert.

Eine wichtige Eigenschaft des Medians ist die Robustheit gegenüber Ausreißern. Man kennt das von seinem Arbeitsweg. Die Erfahrung sagt, dass vom Verlassen der Wohnung bis zur Ankunft am Arbeitsort typisch 30 Minuten benötigt werden, auch wenn der Arbeitsweg wenige Male im Jahr wegen des Wintereinbruchs oder einer Baustelle mehr als eine Stunde beansprucht. Dieser Erfahrungswert ist mit der „Typischen Gleisaufenthaltszeit" in der Analyse der Lokeinsatzzeiten statistisch gesichert. Sie stellt die Sollzeit für den Vergleich mit der operativ ermittelten „Aktuellen Gleisaufenthaltszeit" dar. Ist Letztere größer, wird die Differenz als Standzeit erfasst.

Vergleicht man die Zeit von 38 Minuten auf dem heutigen Arbeitsweg, dann sind die 8 Minuten Verspätung die Abweichung vom typischen Wert. In der Analyse der Lokzeiten ist die (positive) Differenz aus „Aktueller Gleisaufenthaltszeit" und „Typischer Gleisaufenthaltszeit" die Standzeit. Die „Typische Gleisaufenthaltszeit" wird dann als Fahrtzeit zugrunde gelegt. Ist die „Aktuelle Gleisaufenthaltszeit" kleiner oder gleich der „Typischen Aufenthaltszeit", dann ist die Standzeit gleich Null und die „Aktuelle Gleisaufenthaltszeit" gleich der Fahrtzeit.

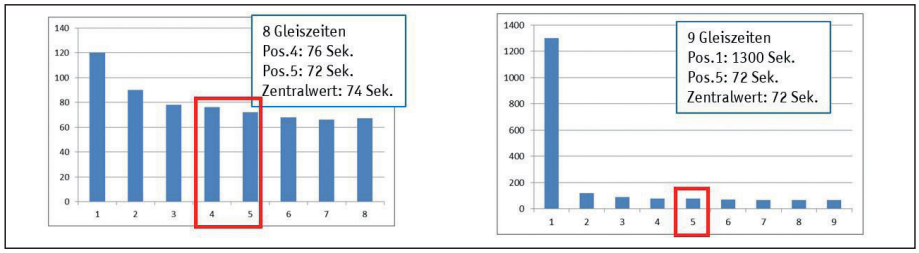

Abb. 2: Ermittlung des statistischen Zentralwerts (Median)

Ladezeit

Für die Ermittlung der Ladezeiten wird davon ausgegangen, dass bei Wechsel des Transportguts der Aufenthalt in Ladegleisen vorrangig der Be- und Entladung dient. Die Durchfahrt durch ein Ladegleis (ohne Wechsel im Feld Transportgut) wird bezüglich Fahrt- und Standzeit identisch zu einem Streckengleis behandelt. In der Stammdatentabelle „Ereignisgleise" wird zusätzlich für jedes Ladegleis eine feste Fahrtzeit definiert. Bei Wechsel des Transportguts wird die statistisch ermittelte „Typische Gleisaufenthaltszeit" abzüglich der fest definierten Fahrtzeit im Ladegleis gleich der Ladezeit gesetzt. Eine Standzeit in der Verladung ergibt sich aus der „Aktuellen Gleisaufenthaltszeit" abzüglich der „Typischen Gleisaufenthaltszeit" (Ladezeit) und der fixen Fahrtzeit.

Abstellzeit

Für Abstellzeiten gibt es zwar manuell erfasste Statuseingaben im Auftragssystem

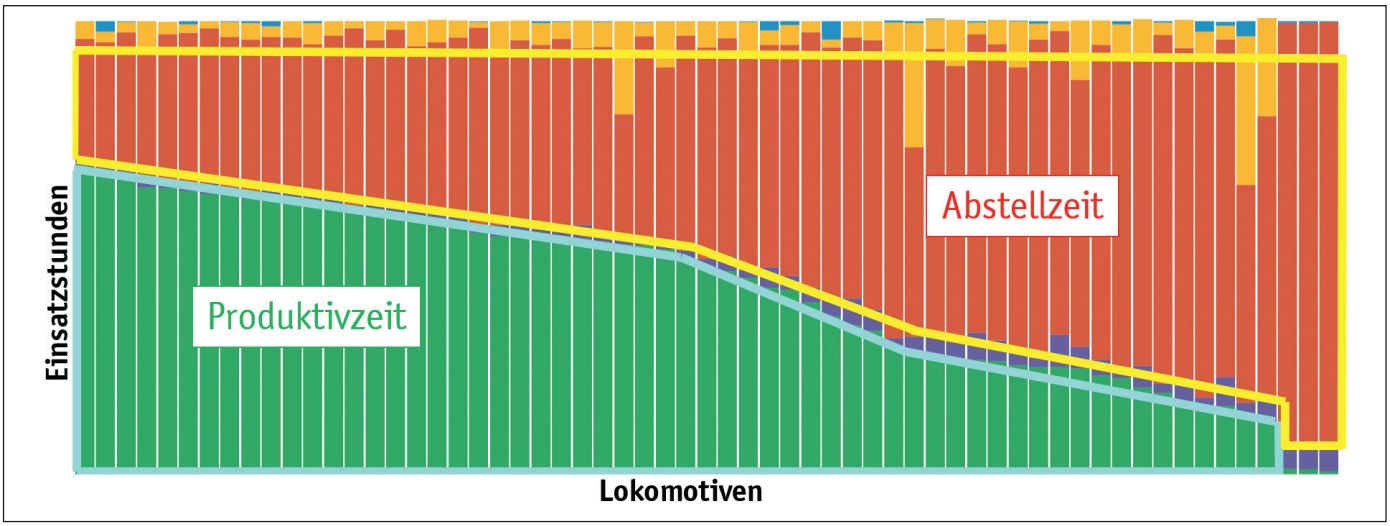

Abb. 3: Analyse Lokeinsatz über ein Jahr

Homepageveröffentlichung unbefristet genehmigt für Zedas GmbH / Rechte für einzelne Downloads und Ausdrucke für Besucher der Seiten genehmigt von DVV Media Group, 2018.

Beschreibung	Fahrzeit Produktiv	Standzeit Produktiv	Be-/Entladezeit	Fahrzeit Nicht Produktiv	Standzeit Nicht Produktiv	Rangierzeit	Abstellzeit Fahrbetrieb	Abstellzeit Werkstatt	Werkstattzeit
Zeitangabe	FZP	SZP	BEZ	FZNP	SZNP	RZ	AZFB	AZWS	WSZ
Formel =	GLFZ-BEZ	AHGLZ-GLFZ	AHGLZ-5Min.	GLFZ	AHGLZ-GLFZ	AHGLZ	AHGLZ	AHGLZ	AHGLZ
Bedingungen									
Einsatzstatus	Im Einsatz	Im Einsatz	Im Einsatz	Im Einsatz	Im Einsatz	Nicht Verfügbar	Verfügbar	Verfügbar	Nicht Verfügbar
Prodstatus	Produktiv	Produktiv	Produktiv	Nicht Produktiv	Nicht Produktiv	Rangiert	Abgestellt	Werkstatt	Werkstatt
weitere Beding.	GLFZ>BEZ	AHGLZ>GLFZ	BelAus/BuAus						

Tab. 2: Berechnungsvorschriften für Zeitenermittlung; GLFZ –Typische Gleisfahrzeit (Median), AHGLZ – Aufenthalts(gleis)zeit, BEZ – Typische Be-/Entladezeit

des Kunden. Sie werden jedoch nicht berücksichtigt, wenn es höher priorisierte, weil technische erfasste, Informationen gibt. So wurde festgelegt, dass eine Lok als fahrbetrieblich abgestellt gilt, wenn es ab einer Stunde keine Gleisfortschaltung für sie gibt. Die Abstellzeit beginnt dann rückwirkend vom Zeitpunkt der letzten Gleisfortschaltung. Die Abstellzeit endet mit der nächsten Gleisfortschaltung.

Rangier- und Werkstattzeit

Das Auftragssystem liefert im beschriebenen Einsatzfall Zeitstempel und Codes, die eine Zuführung zur Werkstatt mit dem Status „Nicht Verfügbar" definieren. Dem entspricht der Status „Rangiert". Die Zeit mit diesem Status fließt in die Rangierzeit ein.

Das Instandhaltungssystem liefert hingegen die Übernahmezeit einer Lok in die Werkstatt (Beginn Werkstattzeit) und die Übergabezeit aus der Werkstatt mit dem Status „Verfügbar" (Ende Werkstattzeit). Das Delta beider Zeiten entspricht der Werkstattzeit. Die Zeit zwischen Verfügbar-Meldung der Werkstatt und der ersten Gleisfortschaltung wird als Abstellzeit-Werkstatt berechnet.

Aufbereitung der Datenquellen

In der Quintessenz entstehen aus den Lösungsansätzen Berechnungsvorschriften (Tab. 2), die auf die zur Verfügung stehenden Informationen aufsetzen.

Bevor sie in Code umgesetzt werden können, sind die technischen Schnittstellen zu den Prozessdatenquellen zu implementieren. Notwendige Zusatzinformationen erfordern den Aufbau statischer Stammdatentabellen und die wiederkehrende Berechnung dynamisch erzeugter Stammdatentabellen. Dazu sind geeignete Softwaresysteme, Computing- und Personalressourcen erforderlich. Laufen die verschiedenen Informationen in den sogenannten Data-Lake ein, werden sie zunächst indiziert und unstrukturiert gespeichert.

Mit Algorithmen zur Datenbasis

Mit Algorithmen (entwickelt für die Analysesoftware oder in Bibliotheken verfügbar) bereiten Datenanalysten die Daten auf, formatieren, filtern und verknüpfen sie zu strukturierten Daten, die sich auswerten lassen. Diese Phase sollte eine Prüfung der Datenqualität ebenso beinhalten wie Plausibilitätsprüfungen. Deswegen sind im Anforderungsprofil der Analysten auch technologische Kenntnisse zu den betrieblichen Prozessen unumgänglich.

Nach Lösung weiterer Detailprobleme wie

- der Zuordnung von mehrtägigen Abstell- oder Werkstattzeiten zu den einzelnen Schichten und Produktionstagen,
- der Berücksichtigung von Schaltjahren, Sommer- und Winterzeit bei der Generierung der Schichtwechselereignisse,
- der korrekten Deduplizierung mehrfach auftretender Eingangsdatensätze und
- einer mehrstufigen Verarbeitung der Daten über den Tag

steht jeden Morgen eine fortgeschriebene Datenbasis zur Erstellung automatisierter Reports bzw. zur manuellen Auswertung bereit. Jedem Datensatz sind die berechneten Zeiten, der Produktions- und Einsatzstatus, zugehörige Informationen aus dem Auftragssystem (Transportgut, Tonnage, Wagenzahl, betriebliche Codes etc.) und dem Instandhaltungssystem (Auftragsnummer, Instandhaltungsstatus unterschieden nach Störung und Frist, Störungskategorie, Fristbezeichner) zugeordnet.

Auswertungen und Nutzung der Daten

Werkzeuge, wie die in diesem Projekt verwendete Splunk-Software, bieten die Aufbereitung komplexer Analyseergebnisse in übersichtlicher und verständlicher Darstellung. Das entbindet nicht von der Aufgabe, sich darüber klar zu werden, welche Fragestellungen beantwortet werden sollen.

Ausgehend vom Ziel, Kostensenkungspotenziale zu finden, wurde zunächst das Verhältnis von Produktivzeit zu Abstellzeit über die gesamte Lokflotte analysiert. Dazu wurden einzelne Zeiten, wie in Tab. 3, dargestellt aggregiert.

Die Jahresauswertung des Lokeinsatzes stellt Abb. 3 dar. Auffällig sind zwei Dinge:

- die ungleichmäßige Verteilung der Produktivzeit über die einzelnen Loks und
- der große Anteil der Abstellzeit.

Einsatzstatus	Produktionsstatus	Zeiten	Zeitkürzel	Zeitaggregierung
Im Einsatz	Produktiv	Fahrzeit Produktiv	FZP	Produktive Einsatzzeit
Im Einsatz	Produktiv	Standzeit Produktiv	SZP	
Im Einsatz	Produktiv	Be-/Entladezeit Produktiv	BEZ	
Im Einsatz	Nicht Produktiv	Fahrzeit Nicht Produktiv	FZNP	Nicht Produktive Einsatzzeit
Im Einsatz	Nicht Produktiv	Standzeit Nicht Produktiv	SZNP	
Verfügbar	Abgestellt	Fahrbetriebliche Abstellzeit	AZFB	Abstellzeit
Verfügbar	Werkstatt	Abstellzeit nach Werkstatt	AZWS	
Nicht Verfügbar	Rangiert	Rangierzeit (Zuführung Werkstatt)	RZ	Werkstattzeit
Nicht Verfügbar	Werkstatt	Werkstattzeit (Geplant und Störung)	WSZ	
Status Nicht Zuordenbar		Ohne Status Zeit	OSZ	Ohne Status Zeit

Tab. 3: Zusammenfassung von Einzelzeiten (Aggregierung)

Homepageveröffentlichung unbefristet genehmigt für Zedas GmbH /
Rechte für einzelne Downloads und Ausdrucke für Besucher der Seiten
genehmigt von DVV Media Group, 2018.

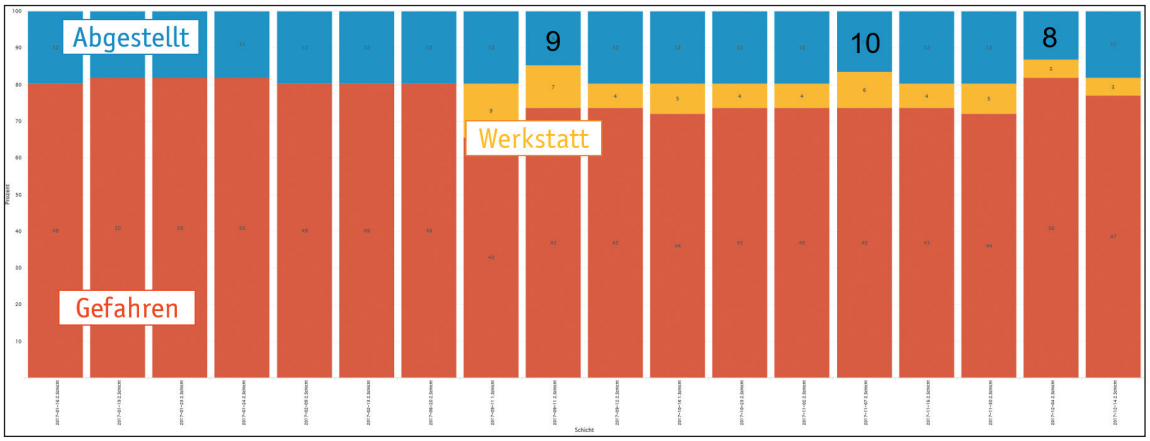

Abb. 4: Schichten eines Jahres mit Unterschreitung der 20 %igen Lokreserve

Eine Schlussfolgerung lautet, dass die betriebliche Verwendung der Loks nicht nach ihrer Belastung erfolgt. Ein Teil der Loks unterliegt über das Jahr einer intensiveren Nutzung und damit einem höheren Verschleiß. Andere Loks etablieren sich als „Standloks". Wenig eingesetzt, werden sie dennoch der fristgerechten Instandhaltung zugeführt.

Eine weitere Schlussfolgerung könnte lauten, dass das Unternehmen zu viele Loks im Bestand hat. Die Produktivzeit lässt sich bei diesem hohen Anteil der Abstellzeit mit weniger Loks erbringen.

Bevor man jedoch „Schnellschüsse" abgibt, sollten weitere betriebliche Randbedingungen analysiert werden. So war zu klären, wie oft eine Spitzenlast bei der Anzahl während einer Schicht eingesetzten Loks auftrat. Auch diese Antwort konnte aus der Datenbasis gezogen werden. In einer Auswertung wurde die Anzahl aller Loks ermittelt, die in der jeweiligen Schicht bewegt wurden, die während der gesamten Schicht abgestellt waren oder die sich während der gesamten Schicht in der Werkstatt befanden. Darauf wurde ein Filter angewandt, der nur die Schichten des vergangenen Jahres einschließt, in denen 20 % oder weniger des gesamten Lokbestandes abgestellt war.

Lediglich in 18 von 1095 Schichten eines Jahres war das der Fall. Es gab drei Schichten, in denen „nur noch" acht, neun oder zehn Loks in Reserve waren (Abb. 4). Nie weniger. Die typische Anzahl (Median) abgestellter Loks lag, je nach Schicht, zwischen 21 bis 32.

Fazit

Der Kunde hat unmittelbar aus der Analyse der Lokeinsatzzeiten die Option zur Senkung seines Lokbestandes in Betracht gezogen; die Umsetzung steht bevor. Weil die Datenaufbereitung und -auswertung als permanenter, automatisierter Prozess aufgesetzt ist, kann er seinen Mitarbeitern in der operativen Lokeinsatzplanung mit der Produktivzeit jeder Lok ein tagesaktuelles Kriterium zur Verfügung stellen, um sie gleichmäßiger auszulasten. Man verspricht sich davon weniger Störungen und Verschleiß, eine effizientere Nutzung der Lokomotiven und damit Kostensenkungen. Mit der Anbindung der Datenanalyse an das Instandhaltungssystem bedarf es zur Erfolgskontrolle lediglich einer weiteren Auswertung. Die notwendige Datenbasis dazu ist vorhanden. Sie wird beim Kunden bereits für eine Reihe weiterer Analysen verwendet und sie bietet viele neue Ansätze. Dazu gehören die Analysen zur örtlichen Verteilung fahrbetrieblicher Standzeiten und zur Relevanz der Einsatzzeit zu auftretenden Lokstörungen.

Grundsätzlich ist das beschriebene System an gleichartige Umgebungen von EVU adaptierbar. Voraussetzung sind technische Schnittstellen zu operativen Prozessdaten, die verlässliche Zeitstempel, Objektidentifikatoren, Ortsangaben und notwendige Zusatzinformationen liefern und ein auf Bahnprozesse spezialisierter „Big Data"-Partner wie etwa die Zedas GmbH. ■

Dipl.-Ing. Ulrich Lieske
Leiter Business Unit Systemintegration / Prokurist
Zedas GmbH, Senftenberg
ulieske@zedas.com