



Quelle: Steag GmbH

Monitoring mit Plus an Intelligenz, Effizienz und Wirtschaftlichkeit

## Prädiktive Instandhaltung auf Basis von Big Data und Machine Learning

Daten aus Energieerzeugungsanlagen in Echtzeit mit unterschiedlichen Sensoren zu erfassen und zu speichern, gehört zum Stand der Technik. Schwankende externe Faktoren wie Wetter, Brennstoffqualität und Last erschweren es jedoch, nur anhand dieser Daten den Zustand der Anlagenkomponenten zu beurteilen. Hierfür müssen die Daten durch intelligente Auswertung in wertvolle Informationen umgewandelt werden. Steag Energy Services ist diesem Ziel mit einem Monitoringsystem basierend auf Big Data und Machine Learning entscheidend nähergekommen.

Die immer größere Leistungsfähigkeit der IT-Hardware ist ein wesentlicher Treiber für Entwicklungen wie Big Data, Digitalisierung und Internet der Dinge. Eine Schlüsseltechnologie der Digitalisierung ist das Machine Learning (maschinelles Lernen). Diese Methode basiert auf Algorithmen, die in den erhobenen Daten bestehende Zusammenhänge automatisch analysieren und hieraus zum Beispiel Vorhersagen treffen (Predictive Analytics) oder wiederkehrende Muster erkennen (Advanced Pattern Recognition).

## Digitale Abbilder durch Machine Learning

Im Umfeld der Energieerzeugung kann das Machine Learning genutzt werden, um digitale Abbilder (Digital Twins) von Anlagen oder deren Komponenten zu erzeugen (**Bild 1**). Machine Learning ist schon seit längerem Bestandteil von Softwarelösungen, die mit datenbasierten, digitalen Abbildern des Prozesses und der Anlagenkomponenten die prädiktive Instandhaltung unterstützen. Aktuelle Entwicklungen im Big-Data-Umfeld umfassen zusätzlich das Potenzial, den Engineering-Aufwand zur Einrichtung einer Lösung für diese Instandhaltungsstrategie zu reduzieren.

## Digitale Twins zeigen relevante Abhängigkeiten

Ein digitaler Twin ist letztlich ein mathematisches Modell, das wiedergibt, wie relevante Prozessparameter wie Leistungen, Temperaturen oder Massenströme von Umgebungsbedingungen und Stellgrößen

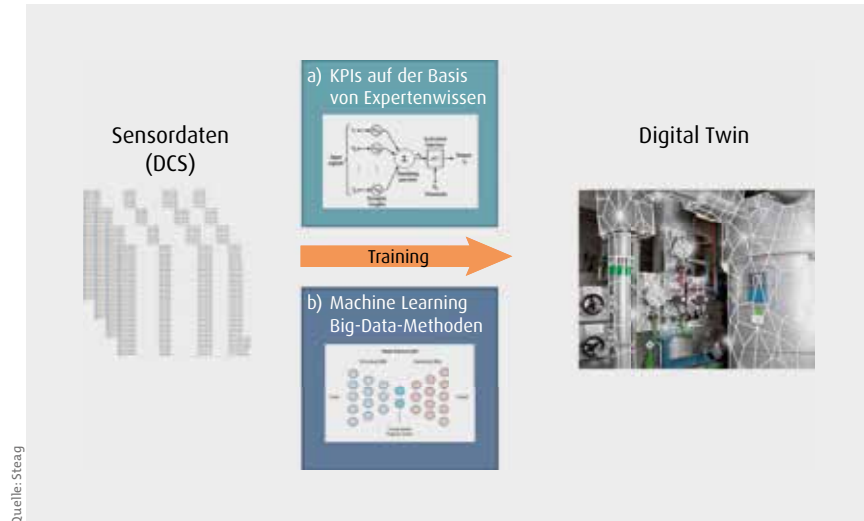


Bild 1. Digital Twins für die prädiktive Instandhaltung

ßen der Anlage abhängen. Das Modell beruht entweder auf physikalischen Grundgleichungen, die die Anlage beschreiben, oder es kann mit Machine Learning aus historischen Daten abgeleitet werden.

In der Folge lassen sich dann die aktuellen Werte betrieblicher Messungen mit den Vorhersagen der Digital Twins vergleichen, damit frühzeitig und belastbar auch schleichende Veränderungen im Anlagenzustand erkannt werden. Auf Basis dieser Informationen kann vorausschauend im Sinne einer prädiktiven Instandhaltung agiert werden, um mit effizientem Ressourceneinsatz Wirkungsgradverluste zu reduzieren und die Anlagenverfügbarkeit zu erhöhen. Die Herausforderung besteht in diesem Zusammenhang allerdings dar-

in, Anomalien im Anlagenverhalten wirklich zuverlässig zu erkennen und gleichzeitig Phantom-Alarme zu vermeiden.

## Mit Expertenwissen zu High Quality KPIs

Zur Erstellung datenbasierter Modelle sind unterschiedliche, aber sich ergänzende Methoden denkbar: die Modellierung durch Kombination von Expertenwissen mit maschinellem Lernen oder durch autonomes, unüberwachtes Lernen (**Bild 2**), weitgehend auf Basis mathematischer Algorithmen mit minimaler ingenieurtechnischer Vorarbeit.

Ein Experte kennt die aussagefähigen Schlüsselgrößen für den Zustand des Pro-

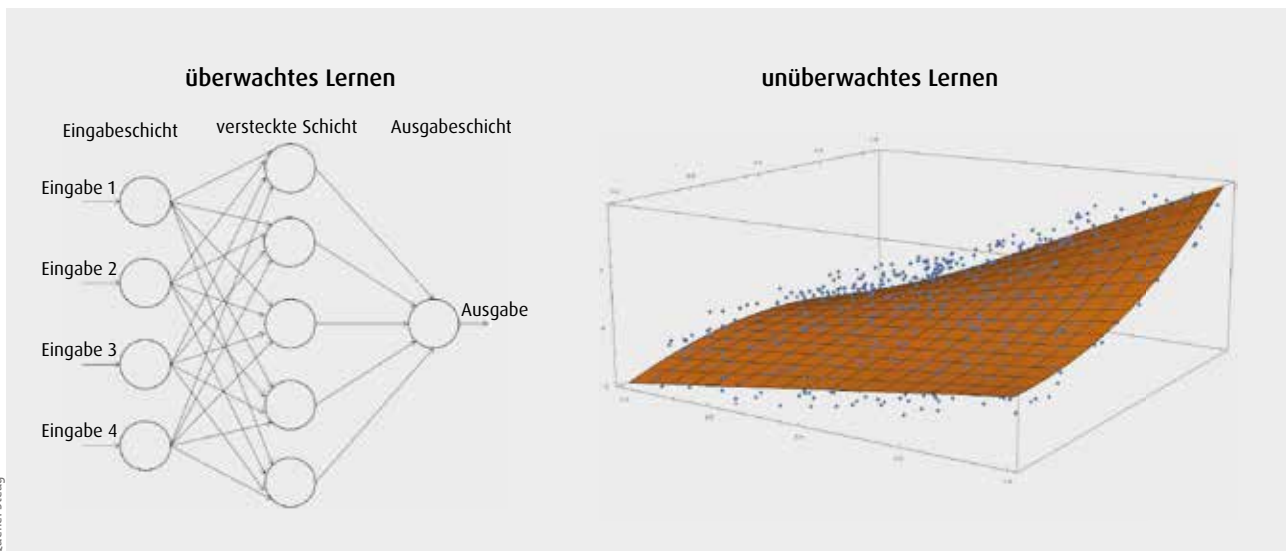


Bild 2. Methoden zur Erstellung datenbasierter Modelle:

Überwachtes Lernen: Expertenwissen über Ursache und Wirkung wird abgebildet. Engineering bei der Modellbildung, aber vereinfachte Auswertung

unüberwachtes Lernen: Korrelationen zwischen den Messwerten werden erkannt und unabhängige Schlüsselgrößen identifiziert.

Automatische Modellbildung, aber ingenieurmäßige Analyse bei der Auswertung

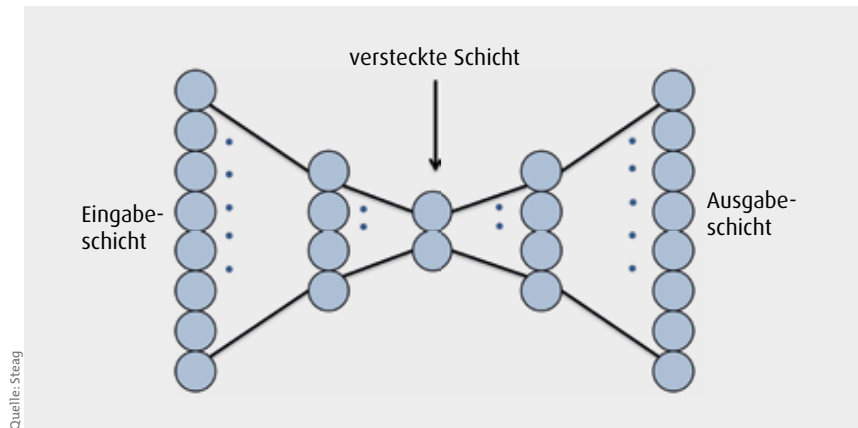


Bild 3. Struktur eines Autoencoders

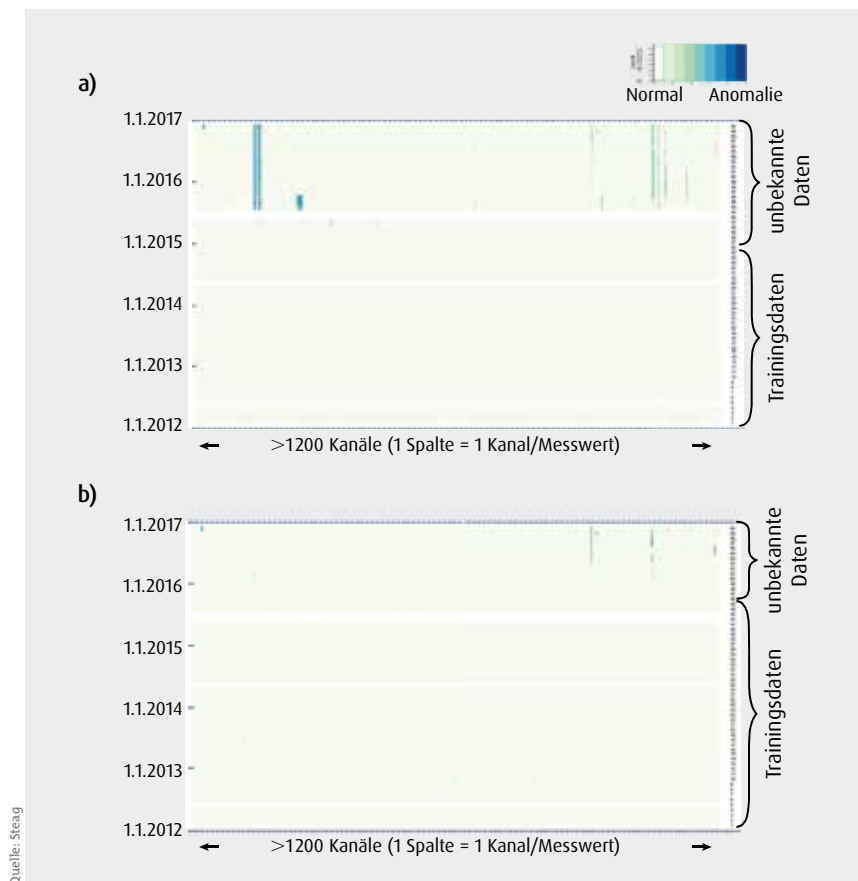


Bild 4. Heatmap der Anomalien in einem Turbosatz

a) Modell mit rund 50 % Lernzeitraum

b) Modell mit rund 75 % Lernzeitraum (einschließlich einer kleinen Teilmenge der Daten nach Revision)

zesses oder eines Anlagenteils. Er weiß auch, welche Einflussgrößen erforderlich sind, um das erwartete Verhalten einer Schlüsselgröße zu beschreiben und welche Zeiträume sich in den historischen Daten als Referenz eignen, also welche Kausalitäten in den Daten vorhanden sind. Der Experte wählt entsprechend die In- und Outputgrößen seines Modells. Überwachte Lernverfahren bilden dann

das Modell, wodurch digitale Abbilder mit hoher Präzision entstehen (High Quality KPIs). Aufgrund der ingenieurtechnischen Auswahl besonders informativer Messwerte, können vom System nachgewiesene signifikante Veränderungen leicht möglichen Störfällen zugeordnet werden, um in der Folge über weitere Untersuchungen oder konkrete Instandhaltungsmaßnahmen zu entscheiden.

### Autoencoder lernt selbstständig

Autonomes Lernen verzichtet hingegen auf die Vorgabe von Kausalitäten in den Daten durch den Experten. Es obliegt dem Machine Learning, bestehende Korrelationen zu identifizieren und entsprechende Modelle zu bilden. Die Algorithmen erkennen dabei, dass die Daten aus einer Anlage nicht unabhängig voneinander sind, sondern von wenigen Schlüsselwerten bestimmt werden. Algorithmen wie der Deep Autoencoder identifizieren daher solche Größen selbstständig und lernen unüberwacht die Zusammenhänge zwischen diesen und den Messwerten einer Anlage.

### Identifizierung wesentlicher Einflussgrößen

Ein Autoencoder (Bild 3) ist ein neuronales Netzwerk, das darauf trainiert ist, Eingabewerte auf Ausgabewerte zu kopieren. Das Netzwerk hat mehrere Schichten verdeckter Lagen von Neuronen und besteht aus zwei Teilen mit einem Flaschenhals in der Mitte. Dieser verhindert, dass das neuronale Netz einfach Eingabe = Ausgabe lernt, und erzwingt die Identifizierung wesentlicher Einflussgrößen (Merkmale), die das Verhalten der Eingangswerte beschreiben.

### Automatische Prüfung signifikanter Abweichungen

In einer Online-Anwendung zur Unterstützung der prädiktiven Instandhaltung werden dem Deep Autoencoder alle zu überwachenden Messungen präsentiert. Beim Lernen identifiziert der Algorithmus die wichtigsten Merkmale, die das Anlagenverhalten beschreiben und wie die Messungen unter normalen Bedingungen hiervon abhängen. Im Online-Modus wird der erste Teil des Autoencoders dann zunächst aus den aktuellen Messwerten die Merkmale bestimmen, die den aktuellen Betriebszustand beschreiben. Der zweite Teil berechnet aus diesen Merkmalen die unter den Betriebsbedingungen zu erwartenden normalen Werte für jede Messgröße. Mit statistischen Methoden wird automatisch für jede Messung geprüft, ob eventuelle Abweichungen zwischen dem aktuellen Wert und der Vorhersage signifikant sind. Auf diese Weise lassen sich auch große Messwertsätze automatisiert überwachen und mit geringem ingenieurtechnischem Aufwand Veränderungen im Zustand der Anlage, die sich in den Messwerten widerspiegeln, sehr frühzeitig und vor allem belastbar erkennen.

Dieser Ansatz, Anomalien in Messwerten zu identifizieren, wurde erfolgreich auf Daten aus verschiedenen Energieerzeugungsanlagen (konventionelle Kraftwerke und Windenergieanlagen) angewandt.

## Anwendungsbeispiel: Erkennung von Anomalien reduziert Brennstoffeinsatz und erhöht Verfügbarkeit

Im Folgenden wird ein Turbosatz eines 1100-MW-Blocks in einem Kohlekraftwerk betrachtet. Dafür liegen mehr als 1200 analoge Kanäle in der Leittechnik vor. Der ausgewertete Datensatz umfasst 5-min-Mittelwerte über 60 Monate, also mehr als 630 Mio. Mittelwerte. Ein Autoencoder, wie oben beschrieben, wurde mit einer Teilmenge der Daten erzeugt. Die Abweichungen zwischen gemessenen und erwarteten Werten wurden für den gesamten Datensatz berechnet. Statistische Analysemethoden identifizieren dann Messwerte, die signifikante Abweichungen (Anomalien) zum normalen Verhalten zeigen.

Die Ergebnisse sind in Heatmaps dargestellt (**Bild 4**). Die Darstellung ist stark komprimiert, um die große Datenmenge überschaubar zu machen. Für jede Messung existiert eine Spalte (insgesamt 1200 Spalten). Jede Zeile stellt einen Zeitabschnitt dar. Der Autoencoder wurde auf 5-min-Daten angewendet. Im Diagramm werden die Ergebnisse auf Monate verdichtet (bei 5 Jahren auf 60 Zeitabschnitte).

### Kombination aus Autoencoder und statistischer Analyse erkennt Anomalien

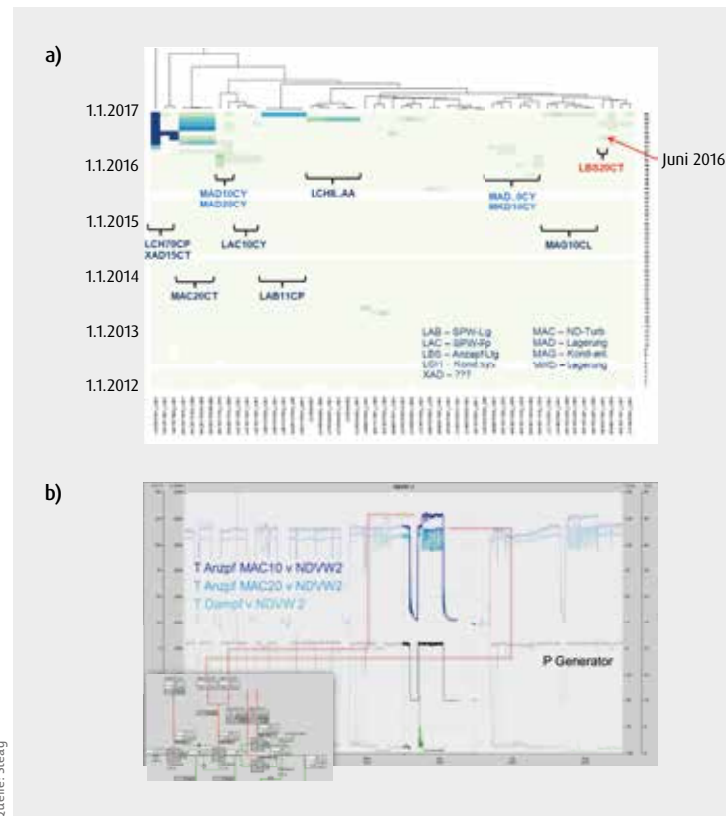
Für jeden Zeitabschnitt haben die Algorithmen gezählt, wie oft die Kombination von Deep Autoencoder und statistischer Analyse eine signifikante Anomalie für die Messung erkannt hat. Die Zeitabschnitte sind in **Bild 4** wie folgt farblich markiert:

- dunkelblau: durchgehend signifikante Anomalien im betrachteten Monat
- hellblau: keine Anomalie in diesem Monat
- weiß: Anlagenstillstand.

Der breite weiße Streifen im oberen Bereich entspricht einer großen Revision des Blocks.

### Anomalien aufgrund von Revisionen ausschließen

Der Autoencoder wurde wie beschrieben mit einer Teilmenge der Betriebsdaten trainiert. Die Trainingsmenge entspricht zunächst etwa den Daten bis kurz vor der Revision. Es ist erkennbar, dass Änderungen im Anlagenverhalten bedingt



**Bild 5.** Veränderungen zuverlässig erkennen  
a) Exklusivansicht der Anomalien in einem Turbosatz (75 % Lernzeitraum)  
b) Detailsansicht einer erkannten Anomalie

durch Maßnahmen im Rahmen der Revision zutreffend als Anomalien ausgewiesen wurden. Um das veränderte Anlagenverhalten nach der Revision abzubilden, wurde der Autoencoder ein zweites Mal mit erweiterten Daten trainiert, die auch Betriebsdaten aus dem Zeitraum nach der Revision enthalten. Der Vergleich der Ergebnisse zeigt, dass Anomalien im späteren Anlagenbetrieb in beiden Fällen erkannt wurden. Anomalien, die durch ein verändertes Anlagenverhalten infolge der Revisionsarbeiten auftraten (**Bild 4a**), wurden durch ein Erweitern der Trainingsdaten als reguläre Daten markiert und entsprechend in der erweiterten Heatmap (**Bild 4b**) wie erwartet nicht mehr ausgewiesen.

### Fehlalarme vermeiden, deutliche Veränderungen zuverlässig identifizieren

**Bild 5** zeigt nur die Messwerte, für die im betrachteten Zeitraum Anomalien aufgetreten sind. Dies war nur für wenige der insgesamt mehr als 1200 Messungen der Fall. Für den Turbosatz einer modernen Anlage ist dies auch so zu erwarten. **Bild 5** verdeutlicht, dass das skizzierte Verfahren in der Lage ist, Fehlalarme zu vermeiden und trotzdem signifikante Veränderungen belastbar zu erkennen. So ließ sich zum Beispiel im Juni 2016 die Anomalie in der Messung LBS20CT001 (rot in **Bild 5a**) einer Fehlfunktion in einer ND-Anzapfung durch

einen defekten Kompensator zuordnen, die auch deutlich im zeitlichen Verlauf der Anzapftemperaturen erkennbar ist (**Bild 5b**).

Das Verfahren ist in der Lage, aus der großen Anzahl von Messgrößen eine überschaubare Anzahl auffälliger Werte belastbar zu identifizieren, wie **Bild 4** und **Bild 5** verdeutlichen. Stichproben haben außerdem gezeigt, dass die ausgewiesenen Anomalien tatsächlich Störungen oder anderen, von der normalen Fahrweise abweichenden Zuständen der Anlage entsprechen.

### Gezielte Vermeidung größerer Schäden

An einem weiteren Block zeigte das System bei einer Nachbetrachtung frühzeitig ein plötzlich auftretendes anomales Laufverhalten der Turbine (**Bild 6**). Genauere Auswertungen ließen einen schweren Turbinenschaden an der MD-Turbine und/oder einen Lagerschaden vermuten. Die Anlage konnte zur Vermeidung noch größerer Schäden sicher abgefahren werden. Das Kaltziehen der Turbine wurde schnellstmöglich eingeleitet, und es wurden Vorbereitungen zur Endoskopie getroffen, die den MD-Turbinenschaden bestätigte.

### Effizienter Einsatz verfügbarer Ressourcen

Das Werkzeug Anomalieerkennung überwacht somit automatisiert ganze



Anlagenteile, wodurch das Betriebspersonal seine Aufmerksamkeit auf diejenigen Bereiche richten kann, in denen sich tatsächliche Veränderungen zeigen. Eine gezielte Analyse ermöglicht es, auf diese Veränderungen mit effizientem Einsatz der verfügbaren Ressourcen im Rahmen einer prädiktiven Instandhaltung vorausschauend zu handeln. Damit werden Verlustquellen frühzeitig beseitigt und ungeplante Stillstände reduziert. Der mittlere spezifische Wärmebedarf wird damit reduziert und die Verfügbarkeit gesteigert.

### Verfügbare Ressourcen effizienter nutzen

Der ökonomische Druck auf Energieerzeugungsanlagen wächst zusehends. Vor dem Hintergrund der insgesamt zunehmenden installierten Leistung wird der wirtschaftliche Betrieb konventioneller Kraftwerke trotz geringerer Auslastung und höherer Belastung der Komponenten durch eine flexiblere Fahrweise immer schwieriger. Für Erneuerbare-Energien-Anlagen, vor allem für Windenergieanlagen, ändern sich

die Fördermodelle. Auch hier wird der gewinnbringende Betrieb von Windparks eine wachsende Herausforderung. Daher ist es entscheidend, frühzeitig und belastbar auch schleichende Veränderungen in einer Anlage zu erkennen, die auf sich anbahnende Störungen oder Ausfälle mit hohen Verlusten hinweisen.

Diese Aufgabe kann selbst erfahrenes Personal nicht leisten, da sie mit zunehmender Variabilität in der Anlagenfahrweise immer schwieriger wird und von immer weniger Menschen geleistet werden muss.

Aus diesem Grunde ist es immens wichtig, die Potenziale wegweisender IT-Entwicklungen wie Big Data und Machine Learning konsequent zu nutzen, um bereits bestehenden Softwarelösungen zur Überwachung von Prozessen und Anlagenkomponenten ein entscheidendes Plus an Intelligenz zu verleihen.

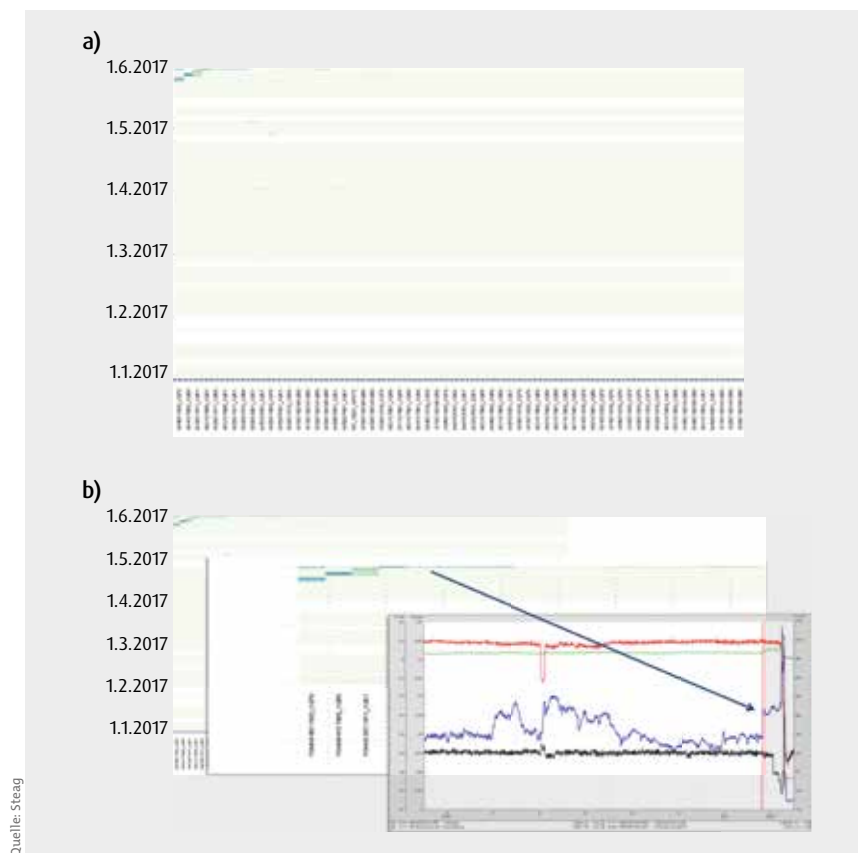
Hieraus ergeben sich leistungsstarke neue Lösungen, die überaus wertvolle Informationen liefern, um signifikante Abweichungen vom regulären Anlagenverhalten ohne hohen personellen Aufwand zeitnah zu identifizieren. Verfügbare Ressourcen lassen sich so wesentlich effizienter einsetzen und Instandhaltungsmaßnahmen noch rechtzeitig und gezielter planen.

>> **Dr. Peter Deeskow**,  
Leiter Produktentwicklung,  
Steag Energy Services GmbH, Essen

**Urs Steinmetz**,  
Produktmanager,  
Steag Energy Services GmbH, Essen

>> [peter.deeskow@steag.com](mailto:peter.deeskow@steag.com)  
[urs.steinmetz@steag.com](mailto:urs.steinmetz@steag.com)

>> [www.steag-systemtechnologies.com](http://www.steag-systemtechnologies.com)



**Bild 6.** Frühzeitiges Erkennen eines anomalen Laufverhalten einer Turbine  
a) Heatmap der Anomalien in einem Turbosatz (1. Halbjahr 2017)  
b) Detailansicht einer frühzeitig erkannten Anomalie im Laufverhalten der Turbine