

# Decision Analytics mit Heatmap-Visualisierung von mehrschrittigen Ensembledaten

## Eine Anwendung von Unsicherheitsmodellierung für Historical Consistent Neural Network und andere Prognosetechniken

Die heutige Leistungsfähigkeit von Rechnern führt leicht zur Generierung von großen Datenmengen. Die wirkliche Herausforderung liegt in einer adäquaten Aggregation der Daten für den Entscheidungsprozess. Vorliegend fokussieren wir auf **Ensemble-Daten, die typischerweise erzeugt werden, wenn Prognose-Verteilungen für mehrere Zeitschritte in die Zukunft erzeugt werden**. Häufig wird eine Verteilung durch Mittelwert oder Median aggregiert. Dies führt zu einem einzelnen Pfad, der leicht zu interpretieren ist. Dieser einzelne Pfad kann jedoch zu ernsthaften Fehlinterpretationen führen, wenn die zugrunde liegende Verteilung sich tatsächlich auf mehrere Pfade aufteilt. Mittelwert und Median können auch in einem Bereich liegen, der nur wenige Ensemble-Mitglieder enthält. Um diesem entgegenzuwirken, schlagen wir vor, Heatmap-Visualisierung zu verwenden, um die Daten für Decision Analytics besser aufzubereiten. Heatmap-Visualisierung ist eine intuitive Möglichkeit, Bereich mit hoher und niedriger Aktivität darzustellen. Die Bereiche werden farbkodiert in Abhängigkeit der (gewichteten) Anzahl von Ensemble-Mitgliedern in einem bestimmten Bereich.

DOI 10.1007/s11576-014-0417-3

### Die Autoren

Dipl.-Math. Cornelius Köpp  
 Prof. Dr. Hans-Jörg  
 von Mettenheim (✉)  
 Prof. Dr. Michael H. Breitner  
 Institut für Wirtschaftsinformatik  
 Leibniz Universität Hannover  
 Königsworther Platz 1  
 30167 Hannover  
 Deutschland  
[mettenheim@iwi.uni-hannover.de](mailto:mettenheim@iwi.uni-hannover.de)

Eingegangen: 2013-07-15  
 Angenommen: 2014-01-31  
 Angenommen nach zwei Überarbeitungen durch die Herausgeber des Schwerpunktthemas.  
 Online publiziert: 2014-04-29

This article is also available in English via <http://www.springerlink.com> and <http://www.bise-journal.org>: Köpp C, von Mettenheim H-J, Breitner MH (2014) Decision Analytics with Heatmap Visualization for Multi-step Ensemble Data. An Application of Uncertainty Modeling to Historical Consistent Neural Network and Other Forecasts. Bus Inf Syst Eng. doi: [10.1007/s12599-014-0326-4](https://doi.org/10.1007/s12599-014-0326-4).

**Zusätzliche Information** ist in der Online-Version dieses Beitrags (doi: [10.1007/s11576-014-0417-3](https://doi.org/10.1007/s11576-014-0417-3)) enthalten.

© Springer Fachmedien Wiesbaden 2014

### 1 Einführung und Motivation

Eine typische Aufgabe für ein Entscheidungsunterstützungssystem ist die Erzeugung von Prognosen und – wichtiger noch – eine Unterstützung bei der In-

terpretation derselben. Punktprognosen können jedoch irreführend sein. Aus diesem Grunde können Ensemble-Prognosen – basierend auf einer Sammlung von Einzelprognosen – verwendet werden, um die Prognosegenauigkeit zu verbessern; siehe z. B. (Zhang und Berardi 2001). Häufig werden auch bei Ensemble-Prognosen nur einzelne Werte (oder einige wenige Werte der Verteilung) verwendet. Dies ist für die automatisierte Verwendung in Informationssystemen nützlich, aber die wirkliche Form der Verteilung bleibt unklar; siehe z. B. (Welch 2001; Hansen 2008). Die Prognose-Verteilung ist nicht notwendigerweise unimodal. Im Allgemeinen kann die Verteilung multimodal sein. Dies bedeutet, dass die Prognose nicht durch einen einzelnen Wert (z. B. den Mittelwert) charakterisiert werden kann. Viele verschiedene Modelle erzeugen die Prognose und es kann nicht erwartet werden, dass die Prognose eingipflig ist. In solchen Fällen können Mittelwert oder Median irreführend sein. Eine Mehrschrittprognose bedeutet, dass nicht nur ein Wert für den kommenden Zeitschritt prognostiziert wird. Vielmehr erstrecken

sich die hier betrachteten Prognosen über mehrere Zeitschritte. Insgesamt können gemeinhin verwendete Aggregationsmethode Informationen verlieren oder skalieren nicht gut. In beiden Fällen wird die Nützlichkeit der Prognose reduziert.

Wir präsentieren erste Schritte zur Beantwortung folgender Forschungsfrage: „Wie ermöglicht eine adäquate Visualisierung Decision Analytics für heutige Ensembleprognosemethoden?“

Ensembles Künstlicher Neuronaler Netze (KNN) sind ein typisches Beispiel für Prognosen die aus mehreren hundert einzelnen Pfaden bestehen. **Im folgenden Beitrag werden 20-Tages-Prognosen des Gaspreises (in US-Dollar) betrachtet.** Die Prognosen werden aus einem Ensemble von 200 Netzen erzeugt. Es wird eine neue Klasse von KNN verwendet: das Historical Consistent Neural Network (HCNN), welches in Zimmermann et al. (2010) eingeführt wurde. Siehe auch von Mettenheim und Breitner (2010) für eine detaillierte Präsentation und Analyse der Performance. Für die Ziele des vorliegenden Beitrags reicht es, sich darüber im Klaren zu sein, dass HCNN eine einfache Zustandsübergangsgleichung verwenden, um den nachfolgenden Zustand aus dem unmittelbar vorgehenden Zustand zu bestimmen. Mehrschrittprognosen sind hierdurch einfach zu generieren. **In der Regel werden HCNN verwendet, wenn mehrere Zeitreihen einschließlich ihrer Verteilung auf einmal modelliert werden sollen.** Werden mehrere HCNNs mit zufällig initialisierten Gewichten trainiert, führt dies zu einem diversifizierten Ensemble.

Die genaue Ausgestaltung der Prognose ist für die folgende Diskussion nicht von zentraler Bedeutung und es ist auch nicht Ziel des Beitrags die Performance dieser KNN-Prognose zu evaluieren. Vielmehr liegt der Fokus des Beitrags darin, eine Prognose für den menschlichen Entscheider leichter interpretierbar zu machen. Der vorliegende Ansatz ist damit nicht auf KNN begrenzt. Andere Modelltypen, die auf Grundlage historischer Daten eine unbegrenzte Anzahl an Prognosen erzeugen, könnten in diesem Kontext ebenfalls verwendet werden. Hierunter fallen z. B. Support Vector Maschinen, Evolutionäre Programmierung oder auch Monte-Carlo-Simulation. Aus diesem Grund wird das KNN Modell nur kurz angerissen. Im Beitrag präsentieren wir eine Heatmap-Visualisierung der resultierenden Ensemble-Prognosen. Dies stellt einen Schritt zu einer visuellen Un-

terstützung des menschlichen Entscheiders dar: Heatmaps aggregieren Informationen, erhalten aber die wesentlichen Prognose-Informationen, auch wenn die Verteilung multimodal ist.

Wir können nun unsere Forschungsfrage konkretisieren: „Wie können wir die vollständigen Prognoseinformationen unter Erhaltung aller Verteilungsinformationen intuitiv für Entscheider darstellen?“

Hierzu schlagen wir eine Heatmap-Visualisierung vor. Eine Heatmap erlaubt es durch Farb-Kodierung, zwischen aktiven und weniger aktiven Bereichen einer Prognose zu differenzieren (s. Abb. 1(d)–1(f) und die Beispiele auf den Abb. 3, 4, 5. Eine detaillierte Diskussion von Abb. 1 folgt weiter unten, insb. in Abschn. 6.

Das einfache Plotten der aggregierten Werte in Abb. 1(a) und 1(b) bewirkt einen Informationsverlust. Es ist auch nicht sinnvoll, schlicht jede Prognose einzeln zu plotten wie dies in Abb. 1(c) geschieht. Einzelne Prognosen können nicht mehr unterschieden werden und das Ergebnis ist nutzlos. **Abbildungen 1(d) und 1(e)** stellen unseren Heatmap-Ansatz vor und **Abb. 1(f)** ist eine sogar noch nützlichere Darstellung. Deutlich können wir sich aufspaltende Pfade in **Abb. 1(d)** und **1(f)** erkennen. Dies ist ein Warnhinweis darauf, dass die Prognose mit einer hohen Unsicherheit behaftet ist.

Heatmaps unterstützen uns auch darin, die Qualität einer Prognose zu beurteilen. In Abhängigkeit der Breite des Prognosekorridors und der Anzahl an Pfaden, die wir in der Heatmap erkennen, können wir die Prognose als eher zuverlässig oder eher unzuverlässig einstufen. Dies ist eine Alternative zum üblichen binären Output mancher Prognosemethoden. Es ist gängig, dass ein Prognosemodell entweder die Prognose „zunehmend“ oder „abnehmend“ abgibt. Dies ist jedoch nicht ganz ehrlich. Es sollte eine dritte Möglichkeit, nämlich „unsicher“, geben. Heatmaps ermöglichen genau diese Art der Ausgabe: anstelle die Information hinter einer einzelnen Zahl, die ohnehin nicht genau treffend sein wird, zu verschleiern, stellen Heatmaps das gesamte Prognosespektrum vor. Problematische Abschnitte lassen sich leicht identifizieren.

## 2 Forschungsdesign

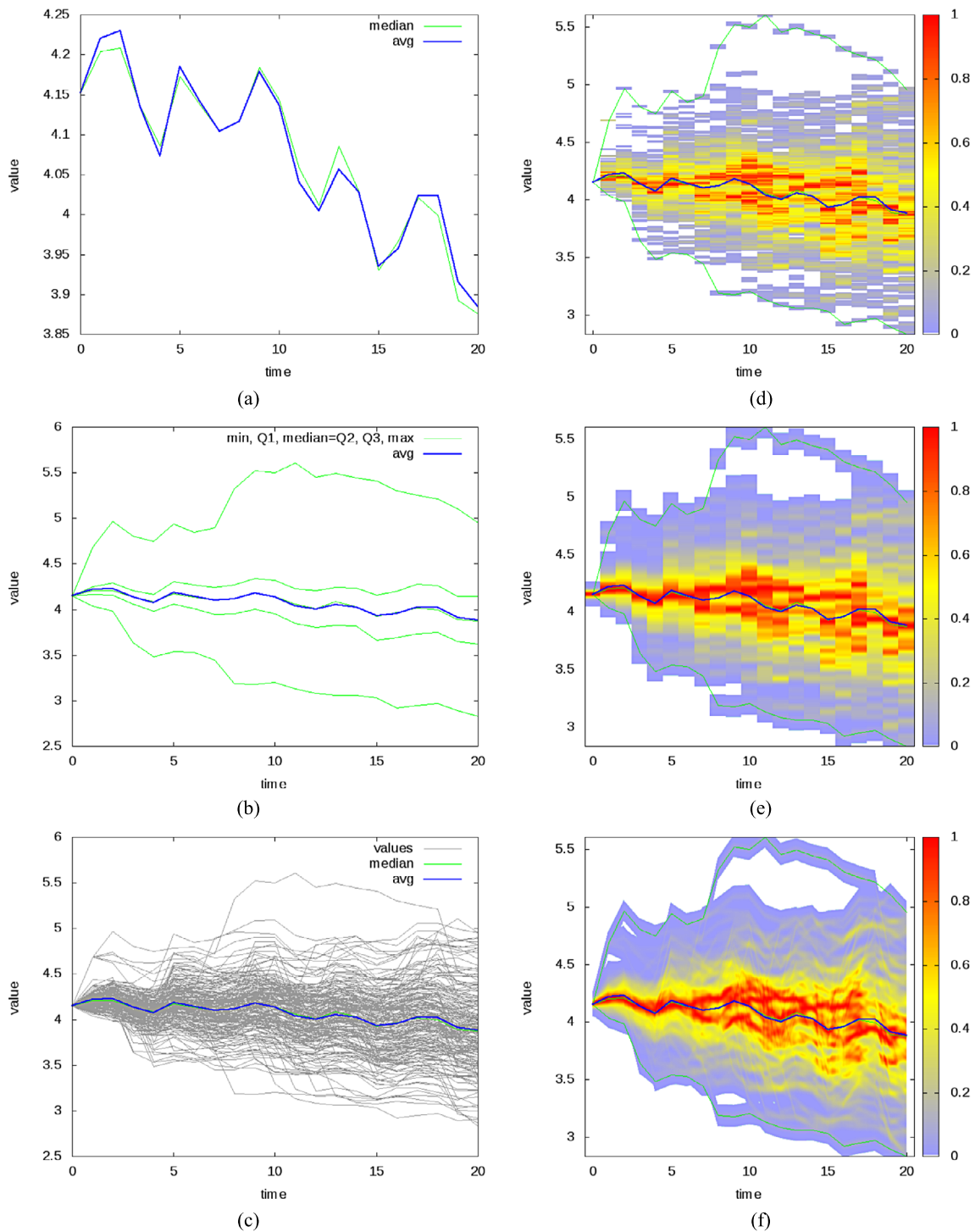
Zur Beantwortung der Forschungsfrage orientieren wir uns am Design-Science-Research-Ansatz von Hevner et al. (2004,

S. 83). **Abbildung 2** zeigt das von uns umgesetzte Forschungsdesign. Im Folgenden stellen wir unseren Beitrag im Kontext ausgewählter Forschungs-Leitlinien dar:

Die vorgestellte Forschung ist relevant (*Problem Relevance*), da unser Visualisierungsansatz Verteilungsprognosen behandelt, die zum Beispiel im Kontext von KNN Ensembles auftreten. Dieselbe Visualisierung könnte nützlich sein für alle anderen Prognosemodell die Ensembles erzeugen, oder Zeitreihen. In Abschn. 3 überprüfen wir existierende Konzepte im Bereich der visuellen Darstellung von Ensembles und Zeitreihen und zeigen die Grenzen gängiger Visualisierungstechniken auf. In Abschn. 4 präsentieren wir das formale Modell und unsere Methode zur Generierung der Heatmaps. Wir argumentieren, dass es sich bei diesem Ansatz um ein Artefakt handelt das „extends the boundaries of human problem solving“ (Hevner et al. 2004) weil es Entscheidungsträgern helfen kann bessere Entscheidungen zu treffen, indem zusätzliche Informationen bereitgestellt werden, die beim Einsatz statischer Aggregation unsichtbar bleiben (*Design as an Artifact*).

Der erste Schritt unserer Evaluation besteht in der Konstruktion eines Prototyps. Dieser wird in Abschn. 5 beschrieben. Wir nutzten den Prototyp zur Demonstration der Nützlichkeit unseres Ansatzes, in Form eines qualitativen Informationsgewinns. Ein Beispielszenario ist durch die oben genannte Gas-Preis-Prognose gegeben, ergänzend dazu dient ein künstliches Szenario zur deutlicheren Darstellung des Informationsgewinns durch Interpolation. Damit folgen wir dem Konzept der Deskriptiven *Design Evaluation* mit Szenarios (Hevner et al. 2004, S. 86).

Ein erster Ansatz des vorgestellten Konzepts wurde bereits Wissenschaftlern und Praktikern auf einer Konferenz (von Mettenheim et al. 2012) und einem Workshop vorgestellt (*Communication of Research*). Wir haben die Ergebnisse mit KNN-Experten diskutiert und deren Feedback berücksichtigt. Auf Basis dieser Diskussion entstand eine unabhängige Implementierung unseres Ansatzes durch erfahrene Anwender in einem großen internationalen Unternehmen. Hevner et al. (2004) fordern, dass „the objective of design-science research is to develop technology-based solutions to important and relevant business problems“. In der Existenz einer unabhängigen Implementierung sehen wir einen



**Abb. 1** Visualisierung von mehrschrittigen Ensembleprognosen

starken Indikator für die Wichtigkeit und Business *Relevance* des durch unseren Ansatz adressierten Problems.

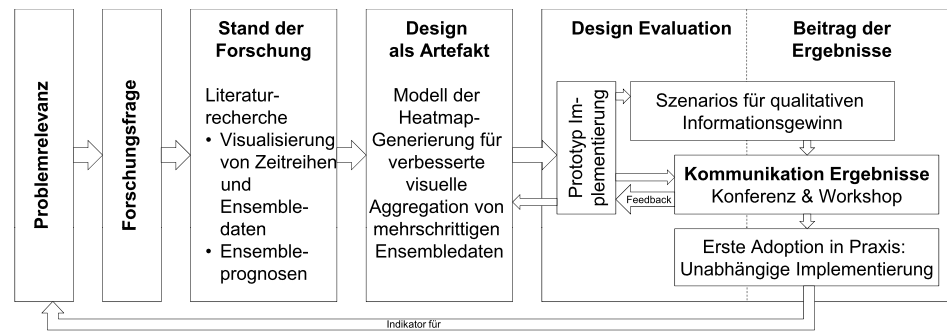
### 3 Stand der Forschung

Potter et al. (2009b) unterstreicht die „enormous power“ von Ensembledaten-

sätzen, aber auch die „formidable challenge“ bei Ensemble-Visualisierung aufgrund ihrer Komplexität. Andrienko und Andrienko (2005) fokussieren sich auf räumlich verteilte Zeitreihendaten wie sie in kartographischen Anwendungen oder Geo-Visualisierung eingesetzt werde. Sie kritisieren das kombinierte Plotten von vielen lokalen Zeitreihen, da überladene

und überlappende Linien zu einer kaum lesbaren Anzeige führen, und das Konzept für eine große Anzahl von hundert und mehr Zeitreihen „completely unusable“ wird. Als Alternative schlagen sie Abbildungs- und Aggregationsbasierte Visualisierung vor. Aggregation wird durch Plotten von Minimum, Maximum, Median und Quartilen reali-

**Abb. 2** Forschungsdesign angelehnt an Hevner et al. (2004, S. 83)



siert. Andrienko et al. (2010) fokussieren sich auf die Unterstützung von Ereigniserkennung in mehreren Zeitreihen. Sie präsentieren ein Toolkit mit interaktiver benutzergesteuerter Datenvisualisierung, mit Mittelwert und Median zur statistischen Zusammenfassung von mehreren Zeitreihen. Bade et al. (2004) nutzt Minima, Maxima, Median und die 25 % und 75 % Perzentile für die Darstellung von aggregierten Hochfrequenzdatenströmen. Hao et al. (2009) beschreibt ein Framework zur visuellen Unterstützung von Zeitreihenprognosen. Ihr Tool nutzt einen ein-dimensionalen heatmapartigen „visual accuracy indicator“ um über-/unter- und nahe Prognosen anzuzeigen.

Aigner et al. (2007) präsentieren ein „conceptual visual analytics framework for time-oriented data“. Sie beschreiben Aspekte der Zeitvisualisierung und zeitorientierten Daten und veranschaulichen einige von denen. Das Untersuchen von Trends und Mustern wird als „particularly important tasks when dealing with time-oriented data and information“ bezeichnet. Sie sehen Interaktion als einen essenziellen Teil von Visualisierung. May et al. (2010) stellen die Definition von Visual Analytics heraus und bezeichnen sie als „the science of analytical reasoning facilitated by interactive visual interfaces“. Thomas und Cook (2006) gehen in dieselbe Richtung. Savikkin et al. (2011) präsentieren ein Visual Analytics Tool für finanzwirtschaftliche Entscheidungsunterstützung.

Buono et al. (2007) präsentieren einen ähnlichkeits-basierten Ansatz der Zeitreihenprognose auf historischen Zeitreihendatensätzen. Ihr interaktives Zeitreihen-Visualisierungs-Tool „Time Searcher 3“ nutzt einen „River Plot View“ und zeigt statistische Merkmale und Quantile zur Angabe der Prognoseunsicherheit. Feng et al. (2010) präsentieren mehrere Methoden zur Visualisierung von Unsicherheit, motiviert durch medizinische An-

wendungen. Sie charakterisieren Dichte-Plots als „useful tools for summarizing extremely large data sets“ und schlagen sie als „fundamental tool for visualizing uncertain multivariate data“ vor. Ihre Visualisierungstechnik für Streudiagramme basiert auf Kerndichtenschätzung (kernel density estimation, KDE) der Wahrscheinlichkeitsdichte in einem Datensatz.

Yagi et al. (2012) präsentieren eine Visualisierungstechnik für das Plotten von getaggten zeitvariablen Daten innerhalb eines einzelnen Linienzugs im Diagrammbereich. Um mit der reduzierten Lesbarkeit von „hunderten oder sogar tausenden“ von Linien umzugehen, nutzen sie einen zweischritten Ansatz: sie clustern die Linienzüge und wählen repräsentative Beispiele für jeden Cluster. Dieses Verfahren haben sie in einer interaktiven Visualisierungsumgebung implementiert.

Zusammenfassend findet die Visualisierung von Zeitreihen-Ensembles durch Plotten von zusammengefassten Daten, vor allem in einfacher Form als Mittelwert/Median/Extrema/Quantile, allgemeine Verwendung; siehe zum Beispiel (Andrienko und Andrienko 2005; Andrienko et al. 2010; Buono et al. 2007; Feng et al. 2010). Der einfachste Ansatz ist die Nutzung von Mittelwert oder Median (s. Abb. 1(a)). Das Hinzufügen von repräsentativen Perzentilen, wie den Quartilen (inklusive Extrema) verbessert die sichtbaren Informationen (s. Abb. 1(b)). Ein weiterer Schritt ist das Plotten aller Pfade der individuellen Ensemblemitglieder; siehe auch Potter et al. (2009a). Abbildung 1(c) zeigt zum Beispiel einen Plot von 200 individuellen Pfaden. Der Ansatz alle Ensemblemitglieder zu plotten skaliert nicht sehr gut. Es wird schwierig, wenn das Ensemble mehrere hundert Mitglieder umfasst. Die „readability of jammed lines is a common problem of information visualization techniques, and several works have addressed the problem“ (Uchida und Itoh 2009).

Ergänzend haben wir mit Google Scholar und Google Images (für den optischen Eindruck) nach ähnlichen Ansätzen gesucht, mit der Kombination aus mindestens zwei der folgenden (englischen) Stichwörter: ensemble; {visualization|visualization}; uncertainty; heatmap; time-series; distribution.

Das Konzept der Fanplots/Fancharts gleicht unserem Ansatz, da eine Heatmap-ähnliche Darstellung für die Visualisierung von Unsicherheit genutzt wird. Fanplots werden zum Beispiel in Makroökonomie und probabilistischen Bevölkerungsprognosen genutzt. Das United Kingdom Monetary Policy Committee (MPC) veröffentlicht Projektionen von BIP und Verbraucherpreisindex in ihrem Inflationsbericht. Elder et al. (2005) dokumentieren diese Nutzung im Detail: die Diagramme basieren auf dem Modus („the single most likely point“) als eine zentrale Projektion mit einer symmetrischen oder schiefen Verteilung. Der farbige Bereich überdeckt nach Definition 90 % des zukünftig realisierten Wertebereichs. Die Breite des Fancharts definiert den Grad der Unsicherheit und steigt mit dem Prognosehorizont, „reflecting the increased probability that some unforeseen event could“ auftreten (Elder et al. 2005, S. 330). Raymer et al. (2012) nutzen Fanplots für die Visualisierung von probabilistischen Bevölkerungsprognosen. Der zweite Autor veröffentlichte das Paket *fanplot* für das R Project of Statistical Computing. Der Plot wird aus den berechneten „percentiles for a set of sequential distributions over a specified time period“ generiert (Abel 2013), wodurch die Möglichkeit existiert Fanplots aus Ensembledaten zu generieren. Eine Einführung und drei beispielhafte Anwendungsfälle werden von Abel (2012) präsentiert. Allerdings ist das Konzept der Fanplots eingeschränkt, da es auf Perzentilen und dem Modus basiert. Die Information von multimodalen Verteilungen und das damit mögliche



**Tab. 1** Formale Symbole und Definitionen

Symbol und Definition	Beschreibung
$m \in \mathbb{N}$	Anzahl der Ensemblemitglieder
$n \in \mathbb{N}$	Anzahl der Prognoseschritte
$F \in \mathbb{R}^{m \times n}$	Ensembleprognose
$M := \{1; \dots; m\} \subset \mathbb{N}$	Indexmenge der Ensemblemitglieder
$N := \{0; \dots; n\} \subset \mathbb{N}_0$	Indexmenge der Prognoseschritte einschließlich Startpunkt
$g \in \mathbb{R}$	Wert des Startpunkts (vor der Prognose)
$D := (gF) = (D_0 \dots D_n) = (d_{ij}) \in \mathbb{R}^{M \times N}$	Eingabedaten (Spalten) für Heatmap-Generierung
$h \in \mathbb{N}$	Prognose-Interpolations-Faktor ( $h = 1$ keine Interpolation)
$N^h := \{0; \frac{1}{h}; \dots; 1; \dots; n - \frac{1}{h}; n\} \subset \mathbb{Q}_+$	Indexmenge der interpolierten Prognoseschritte
$D^h := (d_{ij}), i \in M, j \in N^{*h}$ mit $d_{i,j'+q} := (1-q) \cdot d_{i,j'} + q \cdot d_{i,j'+1}$ ( $j' + q \in N^h, j' \in N$ )	Interpolierte Daten
$x \in X := \{0, \dots, n \cdot h\} \subset \mathbb{N}_0$	Heatmap $x$ -Koordinaten
$y \in Y := \{0, \dots, r\} \subset \mathbb{N}_0$	Heatmap $y$ -Koordinaten
$v_{\min}, v_{\max} \in \mathbb{R}$ mit $d_{ij} \in [v_{\min}, v_{\max}]$	Bereich der in der Heatmap angezeigten Prognosewerte
$V := \{v_y := v_{\min} + y \cdot (v_{\max} - v_{\min}) / r \mid y \in Y\} \subset \mathbb{R}$	Heatmap $y$ -Koordinaten-Werte
$H := (H_0 \dots H_{n-h}) \in \mathbb{R}^{Y \times X}$ mit $H_x := c(D_{x/h}^h) \in \mathbb{R}^Y$	Heatmap-Daten und Heatmap-Spalten-Generator
$c(D_j^h) := (c_j(v_0), \dots, c_j(v_r))^t \in \mathbb{R}^Y$	Heatmap-Spalten-Generator
$c_j(v) := s_j(v) * k(v) = \sum_{i \in M} k(v - d_{ij}) \in \mathbb{R}$	Heatmap-Element-Generierung durch Faltung
$s_j(v) :=  \{i \in M \mid d_{ij} = v\} , j \in N^h$	Anzahl der Ensemble-Mitglieder in Spalte $j$ innerhalb von $v$
$k(\delta) : \mathbb{R} \rightarrow \mathbb{R}$	Faltungskern
$H_{\text{normed}} := (H_0 \dots H_{n-h}) \in \mathbb{R}^{Y \times X}$ mit $H_{\text{normed},x} := H_x / \max(H_x)$	Spaltenweise normierte Heatmap-Daten

che Prognoseergebnis „weiß nicht“ geht verloren.

#### 4 Mathematisches Modell der Heatmap-Generierung

**Tabelle 1** definiert die formale Basis unseres Ansatzes zur Heatmap-Generierung. Im Folgenden werden wir durchgängig den Term Prognosedaten verwenden, obgleich das Modell auf alle geordneten mehrschrittigen Ensembledaten anwendbar ist. Daher müssen die Prozessschritte nicht zwingend Zeit beschreiben.

Zwei Parameter definieren die Dimension der Prognosedaten  $F$ :  $m$  gibt die Anzahl der Ensemblemitglieder zur Erzeugung der Verteilung an. Der Parameter  $n$  gibt die Anzahl der Prognoseschritte an.  $M$  und  $N$  sind die zugehörigen Indexmengen, die benötigt werden, um die Position innerhalb der Prognosedaten zu beschreiben, wobei  $N$  um den Index 0 für den bereits bekannten Startpunkt erweitert wird. Der Wert des Startpunktes ist durch  $g$  gegeben. Die Eingabedaten für die Heatmap-Generierung  $D$  werden aus  $F$  durch voranstellen von Schritt 0 und dortigem Einsetzen des Wertes  $g$  für alle Ensemblemitglieder erzeugt.

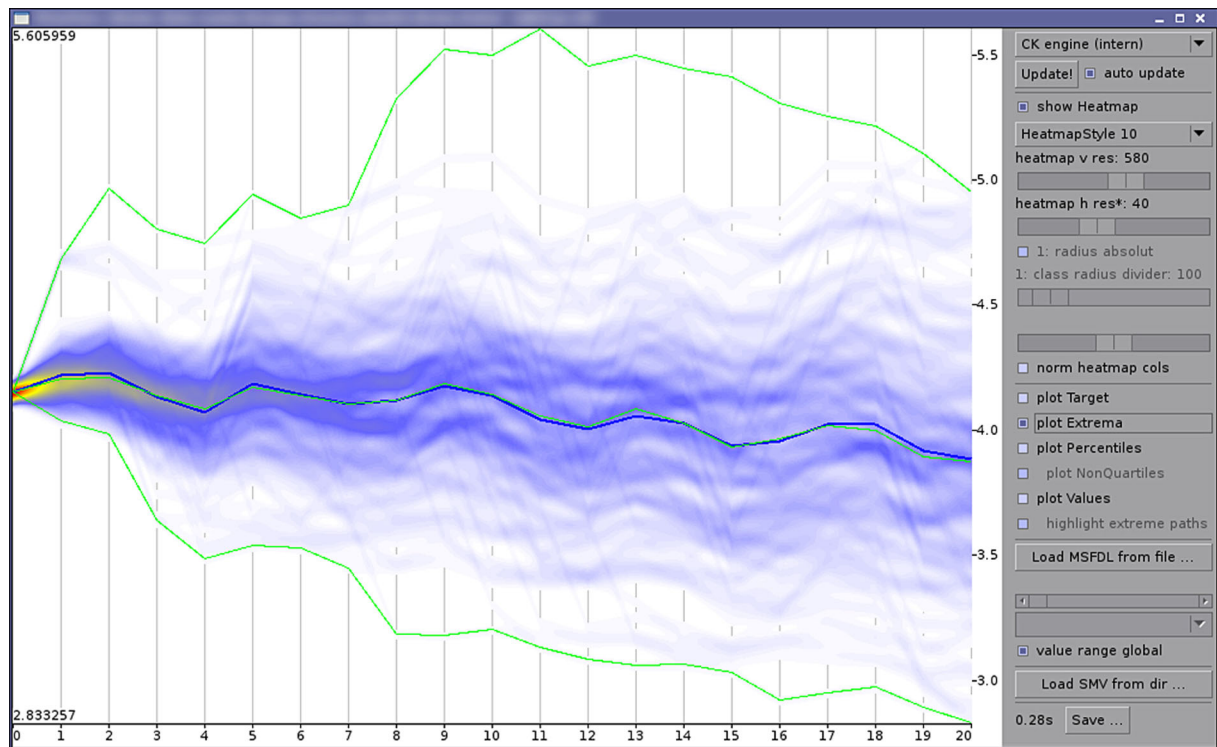
Der Parameter  $h$  beeinflusst direkt die horizontale Auflösung der resultierenden Heatmap. Für  $h > 1$  werden die interpolierten Prognosedaten  $D^h$  durch lineare Interpolation jedes einzelnen Ensemblemitglieds gebildet.  $N^h$  definiert die zugehörige Indexmenge der interpolierten Prognoseschritte. Diese Interpolation erhält die ursprünglichen Prognosedaten und es gilt  $N^{-1} = N$  sowie  $D^{-1} = D$ .

Die Heatmap-Auflösung ist durch die Anzahl der interpolierten Schritte  $n \cdot h$  und die vertikale Auflösung  $y_{\text{res}}$  gegeben.  $X$  und  $Y$  sind die zugehörigen Indexmengen der Heatmap-Koordinaten,  $V$  ist die Menge der den  $y$ -Koordinaten zugeordneten Werte. Der Wertebereich ist definiert als  $[v_{\min}, v_{\max}]$  und sollte mindestens alle Werte aus  $D$  enthalten. Wenn mehrere Heatmaps aus einer Serie von Prognosen generiert werden, so sollte für alle Heatmaps der selbe Wertebereich genutzt werden und alle Werte aller Prognosen enthalten um eine bessere Vergleichbarkeit zu gewährleisten.

Die Heatmap  $H$  wird spaltenweise generiert für jeden Zeitschritt und interpolierten Zeitschritt. Jede Spalte  $H_x$  der Heatmap ist unabhängig von allen anderen Spalten und die Berechnung der Heatmap-Spalten-Generator-Funktion  $c$  hängt ausschließlich von der Spalte  $D_{x/h}^h$

ab. Somit hat die Interpolation keinen Einfluss auf die nicht interpolierten Spalten. Grundsätzlich kann  $c$  jede Funktion des Typs  $c : \mathbb{R}^M \rightarrow \mathbb{R}^Y$  sein, die die Verteilung der Werte in der Daten-Spalte  $D_{x/h}^h$  in ein „nachvollziehbares“ Bild umsetzt. Der Spaltenvektor  $c(D_j^h)$  setzt sie zusammen aus den Berechnungsergebnissen von  $c_j$  für jeden Heatmap- $y$ -Koordinaten-Wert. Für  $c_j$  nutzen wir eine Funktionsklasse die durch eine (diskrete) Faltung eines Kerns  $k(\delta)$  mit  $s_j(v)$  der Anzahl der Ensemblemitglieder an der Stelle  $v$  in Spalte  $k$  nachgebildet werden kann. In den folgenden Abschnitten werden wie verschiedene realisierte Heatmap-Spalten-Generatoren diskutieren.

Die Normalisierung der Heatmap-Spalten ist der optionale letzte Schritt der Heatmap-Berechnung. Durch die Normalisierung kann die Lesbarkeit verbessert werden, da der Startpunkt am Index 0 immer die maximale Dichte aufweist. Diese Dichte ist für gewöhnlich deutlich höher als die Dichte der Prognosen. Durch die Nutzung der Normalisierung ergibt sich eine zusätzliche Anforderung an den Heatmap-Spalten-Generator: es muss sichergestellt sein, dass jeder Spaltenvektor  $H_x$  aus nicht-negativen Werten besteht und mindestens ein Element



**Abb. 3** Screenshot des prototypischen Visualisierungs-Tools (alternative Heatmap-Variante)

mit echt positivem Wert enthält. Dies stellt sicher, dass  $\max(H_x) > 0$  für alle  $x \in X$ , so dass  $H_{\text{normed},x}$  die Division jeder Heatmap-Spalte durch ihr Maximum definiert ist.

Die Werte der resultierenden Heatmap-Matrix  $H$  (oder  $H_{\text{normed}}$  in gleicher Weise) werden farbcodiert dargestellt im Bereich  $[0, \max(H)]$ .

## 5 Prototypisches Visualisierungs-Tool

Das Modell der Heatmap-Visualisierung wurde als Softwareprototyp in Java umgesetzt, mit einer in SWT (Standard Widget Toolkit) realisierten Benutzeroberfläche. **Abbildung 3** zeigt den Prototyp mit einer alternativen Heatmap-Variante und deaktivierter Normierung der Spalten. Insgesamt umfasst der Prototyp 15 verschiedene Heatmap-Spalten-Generatoren, inklusive Algorithmen die auf der Zählung von Ensemble-Mitgliedern innerhalb einer gegebenen  $\varepsilon$ -Umgebung basieren (einfach und sehr schnell), **Faltungen mit Gauß- und anderen ähnlich geformten Kernen, Kern-dichteschätzer**, sowie zwei abweichenden Ansätzen vergleichbar mit einer kumulativen Verteilungsfunktion und ihres

mittleren Gradienten innerhalb des zu einem Heatmap-Pixel zugeordneten Wertebereichs. Die Zählung der Ensemble-Mitglieder innerhalb einer gegebenen  $\varepsilon$ -Umgebung kann ebenso als eine Faltung mit beschrieben  $k_1(\delta) := 1_{\{|\delta| < \varepsilon\}}$  werden.

Im rechten Teil des Programmfensters können verschiedene Parameter der graphischen Darstellung eingestellt werden. Vor allem die Heatmap-Variante und die optionale Normierung, die vertikale Auflösung  $r$  („heatmap v res“) und der horizontale Interpolations-Faktor  $h$  („heatmap h res\*“). Größere Parameter erzeugten glattere Heatmaps, aber erhöhen die für die Heatmap-Berechnung benötigte Zeit. Weitere Bedienelemente erlauben dem User eine Anpassung des in mehreren Algorithmen eingesetzten freien Parameters und des „class radius divider“ („Klassen-Radius-Teiler“) zur Berechnung des Parameters  $\varepsilon$  in  $k_1(\delta)$  durch Division durch  $(\max(H_x) - \min(H_x))$  für eine relative, oder  $(v_{\max} - v_{\min})$  für eine absolute Definition des Radius.

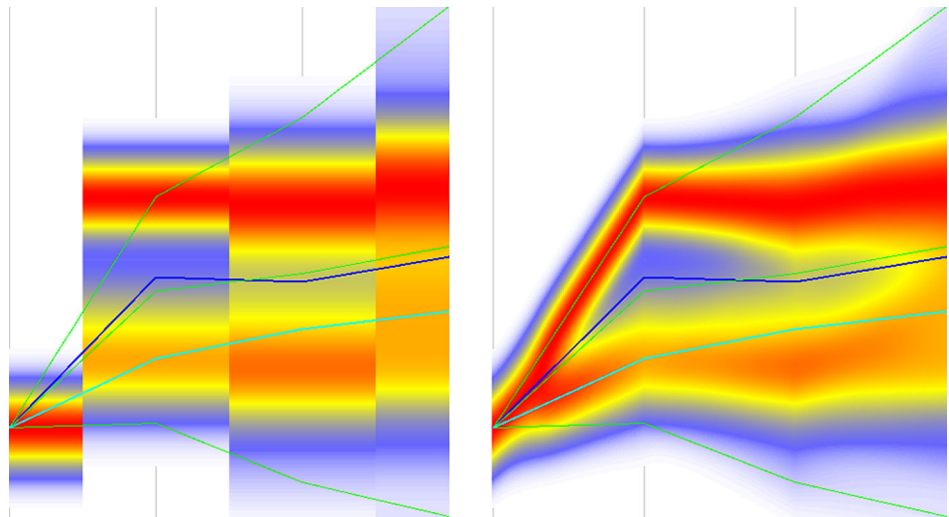
Verschiedene Checkboxes erlauben eine Auswahl der anzuzeigenden Informationen: Extrema, Quantile, Perzentile, alle Ensemblemitglieder mit einer optionalen Hervorhebung aller Mitglieder, die mindestens einmal die Extrema berührt haben, den Zielwert für Analyse der

Prognosequalität und des „Ensembleverhaltens“ aus der Retrospektive, und schließlich auch die Heatmap.

Wir einen besonderen Fokus auf die interaktive Komponente des Prototyps gelegt, da dieser Aspekt von verschiedenen Autoren besonders hervorgehoben wurde (Aigner et al. 2007; Andrienko et al. 2010; May et al. 2010; Potter et al. 2009b). Die erste Plotting Engine mit der freien Software Gnuplot (Version 4.2 Patchlevel 6) erwies sich als zu langsam: hochauflösende Plots benötigten teilweise über 30 Sekunden. Selbst eine direkte Übertragung von binär-kodierten Heatmapdaten in den Eingabestream des Gnuplot-Prozesses führt nicht zu einer ausreichenden Beschleunigung. Die zusätzliche interne Plotting reduziert die Zeit für die Heatmap-Generierung auf unter 0,2 Sekunden – eine Verbesserung um zwei Größenordnungen – und wird nun als Standard verwendet. Die Heatmap-Generierung wird durch die Verteilung der einzelnen Heatmap-Spalten auf die verfügbaren Prozessoren parallelisiert und nutzt somit Mehrkern-CPU's gut aus. Die Berechnung der Heatmap-Spalten wird vereinfacht, zum Beispiel durch Nutzung des Gauß-Kerns  $k_\varepsilon(\delta) := e^{p \cdot \delta \cdot \delta}$  statt  $e^{-\delta^2/(2a^2)}$ . Ein alternativer Algorithmus nutzt eine Lookup-Table (nur einmalig vorberechnet für alle Spalten) um

**Abb. 4**

Informationsgewinn durch  
Zeitschritt-Interpolation:  
Aufspaltende Pfade  
(ohne/mit Interpolation)



die Berechnungsgeschwindigkeit um eine Größenordnung zu verbessern. Dieser Geschwindigkeitsgewinn wurde auf einer üblichen 4-Kern-CPU erreicht. Wir betrachten das verzögerungsfreie Navigieren durch große Prognose-Serien als einen wichtigen Aspekt. Dies ermöglicht den Nutzern zu sehen wie sich die Prognose über die Zeit entwickelt. Darüber hinaus ist eine schnelle Anzeige hilfreich bei Live-Anwendungen. So ist zum Beispiel in Daytrading-Anwendungen die Latenzzeit entscheidend.

## 6 Diskussion

Als ersten Schritt der Ensemblevisualisierung können wir einfach den Mittelwert oder Median plotten (s. Abb. 1(a)). Der allgemeine visuelle Eindruck von Abb. 1(a) ist der eines starken Abwärtstrends. Die Abbildung vermittelt keinerlei Verteilungsinformationen. Dies ändert sich leicht wenn wir repräsentative Perzentile hinzufügen, wie die Quantile (s. Abb. 1(b)). Beachtenswert ist der Umstand, dass zukünftige Unsicherheit – gemessen als die Differenz aus Maximum und Minimum – nicht zwingend ansteigt. Tatsächlich nimmt die Unsicherheit innerhalb der letzten fünf Tage der Prognose ab. Aufgrund der Breite der Verteilung vermittelt die Abbildung nun nur den visuellen Eindruck eines leichten Abwärtstrends. Dennoch haben wir immer noch keine Vorstellung von der Verteilung der individuellen Pfade.

Ein zusätzlicher Schritt ist das Plotten aller 200 Pfade der individuellen Ensemblemitglieder; siehe auch Potter et al. (2009a). Dies führt zu Abb. 1(c). Aus Gründen der Übersichtlichkeit zeigt die

Abbildung auch Mittelwert und Median mit breiten Linien. Es wird offensichtlich, dass durch das Plotten jedes Pfades kein signifikanter Informationsgewinn erfolgt. Ganz im Gegenteil: durch die überlappenden Pfade sind die Informationen weniger klar erkennbar. Sichtbar ist nur dass die Verteilung in der Mitte dicht ist und weniger dicht an den Rändern. Vom visuellen Eindruck her könnte man daher nun schließen, dass die Verteilung unimodal ist – eine möglicherweise gefährliche Schlussfolgerung wie wir später sehen werden.

Rote Regionen zeigen in der Heatmap-Visualisierung eine hohe Aktivität nach der Prognose. Zur klareren Darstellung zeigen die Abb. 1(d)–1(f) sowie Abb. 3 auch Mittelwert, Median, Maximum und Minimum mit breiten Linien. Abbildung 1(e) zeigt einer glattere Heatmap-Variante als Abb. 1(d), und 1(f) zeigt den Stil von Abb. 1(e) geglättet durch Interpolation. Wir halten fest, dass der Mittelwert oft, aber nicht immer, mit den roten Regionen zusammenfällt. Jedoch ist ebenfalls zu erkennen, dass sich die rote Region in Abb. 1(d) und 1(f) ungefähr von Tag 5 bis Tag 15 aufsplittet und sehr groß wird. Dies wird vor allem in der geglätteten Version der Heatmap in Abb. 1(f) deutlich. Daraus lässt sich schließen, dass der Mittelwert laut Prognose die Verteilung nicht sehr gut repräsentiert, weil (speziell an den Tagen 6 bis 8) die Verteilung bi- oder multimodal ist. Dies verändert die Interpretation der Entscheidungsträger signifikant. Betrachten wir nur die Abb. 1(a)–1(c) so scheint es offensichtlich, dass für die Tage 6 bis 8 einen leichter Abwärtstrend prognostiziert wird. Betrachten wir Heatmap so sehen wir, dass die

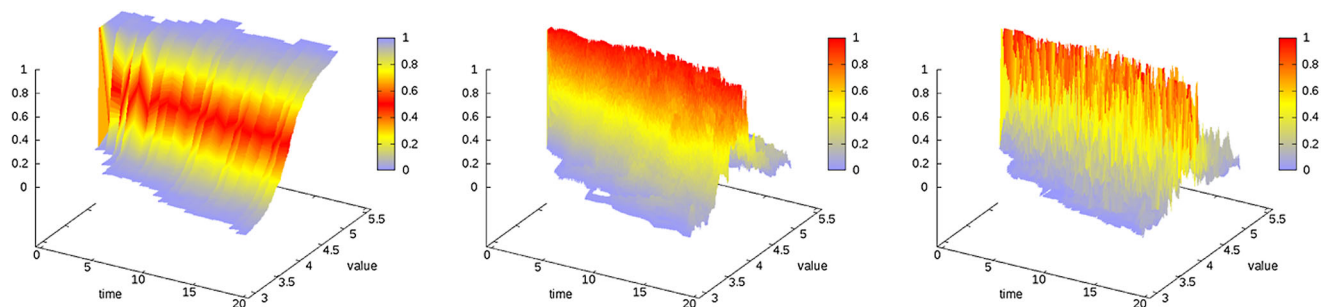
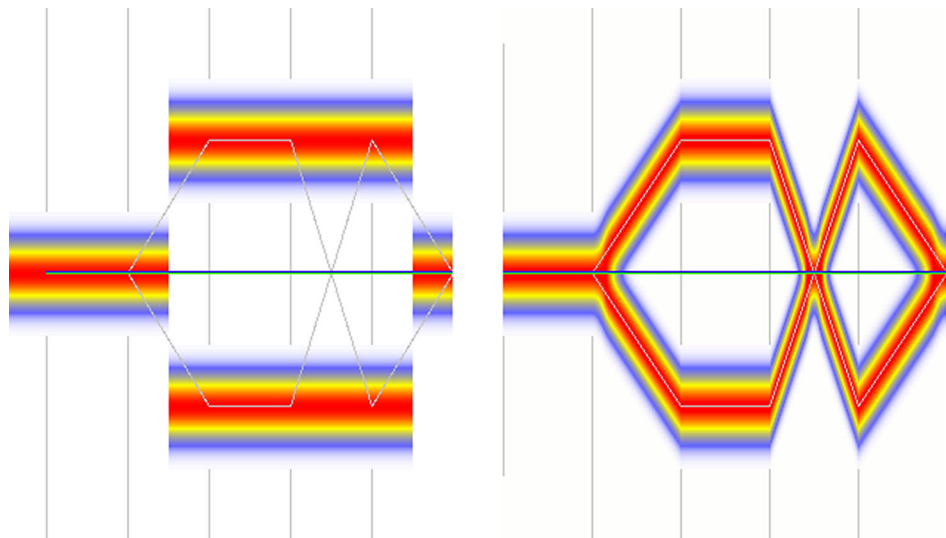
korrekte Antwort lautet: das Ensemble weiß es nicht! Dies ist eine Warnung an die Entscheidungsträger. Andererseits nimmt die Unsicherheit innerhalb der letzten Prognostage (Tag 16–20) ab und das Modell prognostiziert einen klaren Abwärtstrend. Zur Erinnerung: wir beschäftigen uns nicht mit der Prognosegenauigkeit, sondern nutzen lediglich a priori die Prognoseinformationen.

Auf den ersten Blick scheinen durch die Interpolation zwischen verschiedenen Prognoseschritten keine zusätzlichen Informationen gewonnen. Die geglättete Grafik erlaubt uns jedoch den Pfaden in der Prognose zu folgen (s. Abb. 4 und 5 für den Informationsgewinn durch Interpolation). Abbildung 4 zeigt die verbesserte Sichtbarkeit von sich aufspaltenden Pfaden im Prognose-Ensemble und Abb. 5 zeigt ein künstliches Beispiel aus zwei parallelen und sich überkreuzenden Ensemblemitgliedern. Eine Unterscheidung ist nur durch Interpolation möglich (s. Abb. 1(e) ohne und Abb. 1(f) mit Interpolation).

Wir haben 15 Heatmap-Generierungsalgorithmen implementiert. Die Nummerierung spiegelt die Reihenfolge der Implementierung wieder. Einige von diesen liefern optisch ähnliche Ergebnisse, andere erzeugen einen vollkommen anderen optischen Eindruck. Tabelle 1-A im Anhang zeigt einen formalen Überblick über alle implementierten Varianten. Abbildung 1-A im Anhang zeigt das optische Ergebnis unseres Beispiels, jeweils erzeugt mit der Standardkonfiguration. Ein Verändern der Parameter kann dieses Ergebnis signifikant verändern. Generell ist dies erwünscht um die Sichtbarkeit der Verteilung zu optimieren, aber nicht

**Abb. 5**

Informationsgewinn durch  
Zeitschritt-Interpolation:  
Überkreuzende Pfade  
(ohne/mit Interpolation)

**Abb. 6** 3D-Visualisierung zum Vergleich: Die Ergebnisse sind schwieriger zu interpretieren

für alle Parameter. Indikatorfunktionsbasierte Kerne (wie  $k_1$  und  $k_4$ ) sind zwar sehr schnell, haben jedoch den Nachteil, dass sie keine stabilen Ergebnisse liefern bei Variation der vertikalen Heatmap-Auflösung  $r$ . Kontinuierliche (aber nicht zwingend differenzierbare) Faltungskerne sind robuster und vermeiden dieses Problem.

Insgesamt produziert der Gauß-Kern  $k_6$  zufriedenstellende Ergebnisse auf verschiedenen Datensätzen, aber die Berechnung mit diesem Kern ist langsam. Zur Verbesserung der Interaktivität wurden drei ähnlich geformte Kerne  $k_7$ ,  $k_8$  und  $k_9$  implementiert. Kern  $k_7$  erzeugt die ähnlichsten Ergebnisse und ist rund viermal so schnell, die anderen sind langsamer und erzeugen ein suboptimales Ergebnis. Durch die Vorberechnung des Gauß-Kerns und den Einsatz einer Lookup-Table ( $k_{10}$ ), sowie der Reduktion von doppelter auf einfache Fließkommagenauigkeit ( $k_{11}$ ) konnte die Berechnungszeit schließlich um etwa Faktor zehn reduziert werden ohne einen sichtbaren Einfluss. Die Gauß-Kerne mit Lookup-Table sind unabhängig von der Auflösung

der Heatmap, bieten eine gute Sichtbarkeit der Pfade und internen Strukturen gleichzeitig und sind schnell genug für eine interaktive Nutzung.

Die vier implementierten KDE-Kerne ( $k_{11}$ ,  $k_{12}$ ,  $k_{13}$ ,  $k_{14}$ ) haben keine konfigurierbaren Parameter. Dies könnte von Vorteil für die Benutzerfreundlichkeit sein, aber in den von uns getesteten Datensätzen ging die Struktur der Ensemble-Verteilung verloren da die ihre Ausgabe „zu breit“ ist. Kern  $k_3$  unterscheidet sich von den anderen Heatmap-Generatoren, da das Ergebnis einer kumulativen Verteilungsfunktion entspricht. Dies macht ihn zu einer interessanten Wahl, wenn das Unter- oder Überschreiten von Quantilen von Bedeutung ist, und erlaubt eine Visualisierung wie mit Fanplots.

Eine augenscheinliche Alternative zur Heatmap-Visualisierung wäre eine 3D-Darstellung. Diese würde die farbcodierten Informationen durch Höhe ersetzen oder ergänzen. Der Nachteil von drei-dimensionalen Darstellungen besteht jedoch darin, dass häufig einige Teile der Darstellung andere verdecken

(s. **Abb. 6**). Generell ist es schwieriger angemessene Informationen aus einer dreidimensionalen Darstellung zu entnehmen. Die 2D-Heatmap bietet einen guten Kompromiss zwischen Informationsdichte und visueller Interpretierbarkeit.

Das zugrunde liegende Prognosemodell (ein KNN-Beispiel) ist kein zentraler Bestandteil bei der Anwendung unserer Forschungsergebnisse. Die vorgeschlagene Heatmap kann zur Visualisierung aller Arten von Prognosen und überhaupt aller geordneten Daten genutzt werden, wenn diese rund 100 oder mehr Pfade enthalten. Die einzige Annahme besteht darin, dass die Prognose aus einer Verteilung besteht, die nicht durch einige wenige Zahlen charakterisiert werden kann. Dennoch könnten Heatmaps die Einführung von KNN-basierten Prognoseverfahren vereinfachen.

## 7 Limitationen

Auch wenn Heatmap-Visualisierung überzeugende Vorteile bietet, sind einige Einschränkungen diesem Ansatz inhärent. Der größte Nachteil ist, dass visuelle



**Ergebnisse schwer zu quantifizieren sind.** Ein menschlicher Entscheider kann sich aufspaltende Pfade leicht als „unsicheren Zustand“ interpretieren. Es bleibt jedoch die Frage, wie die visuelle Information maschinenlesbar aggregiert werden kann, ohne einen Informationsverlust zu bewirken.

Der Fokus des vorliegenden Beitrags liegt auf der visuellen Darstellung. Es wäre interessant zu untersuchen, ob der Vorteil einer besseren Ausnutzung der Prognoseinformation auch quantifiziert werden kann. Dies bedeutet z. B., dass Gipfel der Verteilung identifiziert werden müssen und ein 3-Zustands-Modell (zunehmend, abnehmend, unsicher) gegenüber einer reinen Prognose auf Basis des Mittelwerts oder Medians gebenchmarkt wird. Andere Verfahren, um die Prognoseinformation besser zu quantifizieren, sind vorstellbar. Schließlich schlagen wir vor, einen Clustering-Algorithmus zur Filterung der Information zu nutzen. Yagi et al. (2012) beschreiben Clustering für Polyline-Plots zeitabhängiger Daten. Idealerweise würde ein solcher Algorithmus dominante Pfade hervorheben und Bereiche mit Rauschen verwerfen.

Eine zweite Kategorie von Nachteilen bezieht sich auf die Anzahl der Meta-Parameters, die zur Visualisierung von Ensemble-Daten gesetzt werden können. **In Abhängigkeit von Parametern wie Kern, Auflösung und Normalisierung der Spalten werden verschiedene visuelle Ausgaben erzeugt. Nicht jede Parameter-Kombination führt zu sinnvollen Ergebnissen in dem Sinne, dass die Interpretierbarkeit der Ergebnisse tatsächlich verbessert wird.** Während andere (Experten-)Nutzer sich nicht über die Komplexität beklagt haben und zügig in der Lage waren, sinnvolle Ergebnisse zu erzeugen, ist die vorliegende Software nicht schlicht „Plug and Play“. Ein grundlegendes Verständnis der verschiedenen Meta-Parameter, insb. die Arbeitsweise von Glättungs-Kernen, ist notwendig. Zurzeit mangelt es an einer automatischen Einstellung der Meta-Parameter.

Die Normalisierung von Spalten ist zweischneidig. Einerseits verbessert Normalisierung die visuelle Darstellung von Verteilungen, die mit fortschreitender Zeit immer breiter werden. Je weiter wir uns in der Darstellung „nach rechts“ bewegen, desto breiter wird in der Regel das Ensemble, was als Unsicherheitsmaß interpretiert werden kann. Es können immer weniger Ensemble-Mitglieder innerhalb eines bestimmten

Radius gefunden werden. Eine Darstellung ohne Normalisierung weist ein erhöhtes Rauschen auf (s. zum Beispiel Abb. 3). Allerdings verschleiert Normalisierung die Information, dass immer weniger Ensemble-Mitglieder zu einem bestimmten Pfad beitragen. Ein menschlicher Entscheider könnte irrtümlich zu dem Schluss gelangen, dass sich der Pfad aus vielen Ensemble-Mitgliedern zusammensetzt wenn es sich doch eigentlich nur um eine kleine Mehrheit handelt. Zusammenfassend sollte die Normalisierungsoption des Prototyps mit Vorsicht verwendet werden. Ein horizontaler Indikator, ähnlich dem von Hao et al. (2009, S. 1) verwendeten, oder eine zusätzliche Kodierung durch die Sättigung könnte die Gefahr einer Fehlinterpretation senken.

Eine potentielle Verbesserung menschlicher Entscheidungen durch Heatmap-Visualisierung wurde bereits mit Experten diskutiert, jedoch noch nicht validiert. Das Feedback von Forschern und Praktikern anlässlich von Konferenz- und Workshop-Präsentationen und in praktischen Anwendungen ist positiv. Eine empirische Validierung der (wahrgenommenen) Nützlichkeit würde jedoch formale Methoden erfordern, z. B. die Verwendung von Experten- und Nutzer-Fragebögen. Selbst wenn ein positiver Effekt von Entscheidern wahrgenommen wird bedeutet dies noch nicht, dass auch die Entscheidungen selbst tatsächlich besser ausfallen. Es könnte schwierig sein, den Beitrag einer besseren Daten-Visualisierung in gemeinhin komplexen Entscheidungsprozessen zu beziffern; siehe hierzu auch Potter et al. (2012, S. 239), die die Evaluation von Unsicherheits-Visualisierung als „double problematic“ bezeichnen. Im speziellen Fall von kurz- und mittelfristigem Handel haben bessere Entscheidungen eine unmittelbare ökonomische Auswertung. Die Nützlichkeit der Methode könnte dadurch untersucht werden, dass Händlern Zugang zu Heatmap-Visualisierung gewährt wird. Aber auch dann bleibt die Schwierigkeit, zwischen der Qualität der Prognosemethode an sich und dem Zusatznutzen durch bessere Visualisierung zu differenzieren. In diesem Kontext könnte sich schließlich eine rein quantitative Validierung ohne menschliche Entscheider doch als überlegen erweisen. Dies würde einen oben erwähnten Clustering-Algorithmus erfordern, der wichtige Pfade automatisch hervorhebt.

## 8 Fazit und Ausblick

Unser Paper zeigt Schritte in Richtung einer adäquaten Visualisierung und Interpretation von Ensemble-Prognosen und verteilten geordneten Daten. Heatmaps bieten einen besseren Weg als viele gängige Methoden, um Prognose-Daten zu nutzen. Wir erkennen, dass Mittelwert und Median nicht notwendigerweise die richtige Information wiedergeben, weil sich zum Beispiel die Verteilung aufspaltet. Dies führt dazu, dass Mittelwert und Median schlechte Repräsentanten der Ensemble-Werte werden. Die Prognose-Unsicherheit (gemessen anhand der Breite des Ensembles) nimmt mit weiteren Zeitschritten nicht einförmig zu. Eine gesplante Verteilung kann wieder eingipflig werden. Heatmap-Visualisierung bietet eine dritte Alternative für menschliche Entscheider: zu den Ergebnissen „zunehmend“ und „abnehmend“ ergänzen wir einen Zustand der Modellunsicherheit „weiß nicht“. Auf diese Weise ergibt sich eine ehrlichere Prognose, da das Risiko einer Fehlinterpretation aufgrund einer einzelnen Zahl reduziert wird. Vorliegend akzeptieren wir explizit die Unsicherheit von Prognosen und verlassen uns nicht auf eine einzelne Zahl.

Der Prototyp bietet verschiedene Heatmap-Varianten, die hier nur kurz diskutiert wurden. Wir benötigen weitere Untersuchungen in Bezug auf die Auswahl passender Heatmap-Varianten und Parametern, da Parametrierung „vital for ensuring effective visualization“ (Aigner et al. 2007) ist. Diesbezüglich fügt unser Ansatz Interpolation zwischen Zeitschritten hinzu, um einen besseren und glatteren Eindruck der zugrundeliegenden Pfade zu erhalten.

Der realisierte Pfad der prognostizierten Zeitreihe wird nicht gezeigt, da unser Fokus auf der Einführung von Heatmap-Visualisierung liegt. In zukünftigen Arbeiten ist es wichtig auch tatsächlich die Performance einer Heatmap-Visualisierung zu benchmarken: wir müssen die Frage beantworten, ob die zusätzliche Information tatsächlich beim Füllen besserer Entscheidungen hilft. Benchmarking ist jedoch eng mit dem zugrundeliegenden Prognose-Modell verbunden; in diesem Falle KNN. Ein Benchmarking hätte den Fokus des Beitrags verschoben, hin zur Prognosequalität von KNN und insbesondere HCNN.

In zukünftigen Arbeiten ist geplant, eine entsprechende Fallstudie vorzustel-

## Zusammenfassung / Abstract

Cornelius Köpp, Hans-Jörg von Mettenheim, Michael H. Breitner

### Decision Analytics mit Heatmap-Visualisierung von mehrschrittigen Ensembledaten

#### Eine Anwendung von Unsicherheitsmodellierung für Historical Consistent Neural Network und andere Prognosetechniken

Heutige in verschiedenen Informationssystemen integrierte Prognosetechniken nutzen oft Ensembles zur Darstellung verschiedener zukünftiger Szenarien. Die Aggregation dieser Prognosen stellt eine anspruchsvolle Aufgabe da: Bei der Nutzung von Mittelwert und Median (gängige Praxis) gehen wichtige Informationen verloren, vor allem wenn die zugrunde liegende Verteilung zu jedem Schritt multimodal ist. Um dies zu vermeiden präsentieren wir einen Heatmap-Visualisierungsansatz. Visuell ist eine einfache Unterscheidung zwischen Bereichen mit hoher Aktivität (hohe Wahrscheinlichkeit der Realisierung) und solchen mit niedriger Aktivität möglich. Diese Form der Darstellung ermöglicht eine Identifikation von sich aufspaltenden Pfaden im Prognoseensemble und schafft dadurch eine „dritte Alternative“ im Entscheidungsraum. Die meisten Prognosensysteme bieten nur Ergebnisse „auf“ oder „ab“ an. Die vorgestellte Heatmap-Visualisierung führt zusätzlich ein Ergebnis „weiß nicht“ ein. Durch Blick auf die Heatmap können somit Bereiche identifiziert werden, in denen sich das zugrunde liegende Prognosemodell nicht sicher ist über den zukünftigen Ausgang. Wir präsentieren einen Softwareprototyp zur Unterstützung von Entscheidern durch eine interaktive Visualisierung und diskutieren den Informationsgewinn durch die Nutzung. Der Prototyp wurde bereits anderen Forschern und Praktikern präsentiert und mit diesen diskutiert.

**Schlüsselwörter:** Decision Analytics, Prognose, Visualisierung, Fuzzy Decision Making, Unsicherheitsmodellierung, Historical Consistent Neural Network (HCNN), Heatmap

### Decision Analytics with Heatmap-Visualization for Multi-step Ensemble Data

#### An Application of Uncertainty Modeling to Historical Consistent Neural Network and Other Forecasts

Today's forecasting techniques, which are integrated into several information systems, often use ensembles that represent different scenarios. Aggregating these forecasts is a challenging task: when using the mean or median (common practice), important information is lost, especially if the underlying distribution at every step is multimodal. To avoid this, the authors present a heatmap visualization approach. It is easy to visually distinguish regions of high activity (high probability of realization) from regions of low activity. This form of visualization allows to identify splitting paths in the forecast ensemble and adds a "third alternative" to the decision space. Most forecast systems only offer "up" or "down": the presented heatmap visualization additionally introduces "don't know". Looking at the heatmap, regions can be identified in which the underlying forecast model cannot predict the outcome. The authors present a software prototype with interactive visualization to support decision makers and discuss the information gained by its use. The prototype has already been presented to and discussed with researchers and practitioners.

**Keywords:** Decision analytics, Forecast, Visualization, Fuzzy decision making, Uncertainty modeling, Historical consistent neural network (HCNN), Heatmap

len, in denen die Ergebnisse der Verwendung von Decision Analytics auf der Grundlage von Heatmaps im Rahmen eines Investment-Prozesses untersucht werden. Ein weiterer interessanter Forschungsbereich ist die Analyse der Auswirkung von Heatmap-Visualisierung mit einem Technology-Acceptance oder Information-System-Success-Modell. Eine unabhängige Implementierung durch eine deutsche Forschungs- und Beratungsgruppe zeigt, dass der Zugriff auf den Informationsgehalt einer Ensemble-Verteilung auch für andere vielversprechend aussieht.

## Literatur

- Abel GJ (2012) The fanplot package for R. <http://gjabel.wordpress.com/2012/08/13/the-fanplot-package-for-r/>. Abruf am 2013-10-21
- Abel GJ (2013) Fanplot: visualisation of sequential probability distributions using fan charts. <http://cran.r-project.org/web/packages/fanplot/index.html>. Abruf am 2013-10-21
- Aigner W, Bertone A, Miksch S, Tominski C, Schumann H (2007) Towards a conceptual framework for visual analytics of time and time-oriented data. In: Henderson SG, Biller B, Hsieh MH, Shortle J, Tew JD, Barton RR (Hrsg) Proc 2007 winter simulation conference, Washington, S 721–729
- Andrienko G, Andrienko N (2005) Visual exploration of the spatial distribution of temporal behaviors. In: Proc ninth international conference on information visualisation. IEEE Computer Society, London, S 799–806
- Andrienko G, Andrienko N, Mladenov M, Mock M, Poelitz C (2010) Extracting events from spatial time series. In: Proc 14th international conference on information visualisation. IEEE Computer Society, London, S 48–53
- Bade R, Schlechtweg S, Miksch S (2004) Connecting time-oriented data and information to a coherent interactive visualization. In: Proc of ACM conference on human factors in computing systems (CHI'04). ACM, New York, S 105–112
- Buono P, Plaisant C, Simeone A, Aris A, Shneiderman B, Shmueli G, Jank W (2007) Similarity-based forecasting with simultaneous previews: a river plot interface for time series forecasting. In: Proc 11th international conference information visualization (2007). IEEE Computer Society, Zurich. doi:10.1109/IV.2007.101
- Elder R, Kaperanios G, Taylor T, Yates T (2005) Assessing the MPC's fan charts. Bank of England quarterly bulletin: autumn 2005. Bank of England, London, S 326–348
- Feng D, Kwock L, Lee Y, Taylor RM II (2010) Matching visual saliency to confidence in plots of uncertain data. IEEE Transactions on Visualization and Computer Graphics 16(6):980–989
- Hansen BE (2008) Least-squares forecast averaging. Journal of Econometrics 146(2):342–350. doi:10.1016/j.jeconom.2008.08.022
- Hao MC, Janetzko H, Sharma RK, Dayal U, Keim DA, Castellanos M (2009) Poster: visual prediction of time series. In: Proc IEEE

- symposium on visual analytics science and technology (VAST 2009), S 229–230
- Hevner AR, March ST, Park J, Ram S (2004) Design science in information systems research. *MIS Quarterly* 28(1):75–105
- May R, Hanrahan P, Keim DA, Shneiderman B, Card S (2010) The state of visual analytics: views on what visual analytics is and where it is going. In: *Proc IEEE symposium on visual analytics science and technology (VAST)*, Salt Lake City, S 257–259
- Potter K, Wilson A, Bremer PT, Williams D, Doutriaux C, Pascucci V, Johnson CR (2009a) Ensemble-vis: a framework for the statistical visualization of ensemble data. In: *Proc international conference on data mining workshops*, S 233–240
- Potter K, Wilson A, Bremer PT, Williams D, Pascucci V, Johnson C (2009b) A flexible approach for the statistical visualization of ensemble data. In: *Proc IEEE ICDM workshop on knowledge discovery from climate data*, Miami
- Potter K, Rosen P, Johnson CR (2012) From quantification to visualization: a taxonomy of uncertainty visualization approaches. In: *Dienstfrey AM, Boisvert RF (Hrsg) Uncertainty quantification in scientific computing*. IFIP advances in information and communication technology, vol 377, S 226–249. doi:[10.1007/978-3-642-32677-6\\_15](https://doi.org/10.1007/978-3-642-32677-6_15)
- Raymer J, Abel GJ, Rogers A (2012) Does specification matter? Experiments with simple multiregional probabilistic population projections. *Environment and Planning A* 44(11):2664–2686. doi:[10.1068/a4533](https://doi.org/10.1068/a4533)
- Savikhin A, Lam HC, Fisher B, Ebert DS (2011) An experimental study of financial portfolio selection with visual analytics for decision support. In: *Proc 44th Hawaii international conference on system sciences (HICSS)*. doi:[10.1109/HICSS.2011.54](https://doi.org/10.1109/HICSS.2011.54)
- Thomas JJ, Cook KA (2006) A visual analytics agenda. *Computer Graphics and Applications*, IEEE 46(1):10–13
- Uchida Y, Itoh T (2009) A visualization and level-of-detail control technique for large scale time series data. In: *Proc 13th international conference information visualisation*, Barcelona, S 80–85. doi:[10.1109/IV.2009.33](https://doi.org/10.1109/IV.2009.33)
- von Mettenheim HJ, Breitner MH (2010) Robust decision support systems with matrix forecasts and shared layer perceptrons for finance and other applications. In: *Proc ICIS 2010*, St. Louis, Paper 83
- von Mettenheim HJ, Köpp C, Breitner MH (2012) Visualizing forecasts of neural network ensembles. In: *Klatte D, Lüthi HJ, Schmedders K (Hrsg) Operations research proceedings 2011*, Zurich, S 573–578. doi:[10.1007/978-3-642-29210-1\\_91](https://doi.org/10.1007/978-3-642-29210-1_91)
- Welch I (2001) The equity premium consensus forecast revisited. *Cowles Foundation discussion paper No 1325*, University of California, Los Angeles (UCLA), National Bureau of Economic Research (NBER)
- Yagi S, Uchida Y, Itoh T (2012) A polyline-based visualization technique for tagged time-varying data. In: *Proc 16th international conference on information visualisation*, Montpellier, S 106–111. doi:[10.1109/IV.2012.28](https://doi.org/10.1109/IV.2012.28)
- Zhang GP, Berardi VL (2001) Time series forecasting with neural network ensembles: an application for exchange rate prediction. *The Journal of the Operational Research Society* 52(6):652–664
- Zimmermann HG, Grothmann R, Tietz C, von Jouanne-Diedrich H (2010) Market modeling, forecasting and risk analysis with historical consistent neural networks. In: *Selected papers of the annual international conference of the German Operations Research Society*, Munich, S 531–536