

# Qualitätsbewertung von Korrekturmethode für hydrographische Zeitreihen

Emanuel ROBATSCH, Alfred WIESER und Karl-Heinrich ANDERS

Fachhochschule Kärnten, Villach · emanuel.robatsch@edu.fh-kaernten.ac.at

*Dieser Beitrag wurde durch das Programmkomitee als „reviewed paper“ angenommen.*

## Zusammenfassung

Das Hochwasser im Oktober 2012 hat gezeigt, dass Hochwasserereignisse auch in Kärnten vorkommen können. Bei der Vorhersage von Hochwassern spielt besonders die Qualität der gemessenen hydrographischen Daten eine wichtige Rolle. Jedoch treten wie bei allen Messungen einzelne Fehlmessungen auf und können somit das Gesamtergebnis, auf dem das Vorhersagemodell basiert, verfälschen. Etwaige Fehler sollten somit in kürzester Zeit aufgespürt und wenn möglich so korrigiert werden, dass sie das Gesamtergebnis nur minimal verfälschen. Die Prüfung und Korrektur der einzelnen Zeitreihen passiert mithilfe von vordefinierten Prüf- und Korrekturmethode. Ziel dieser Arbeit es ein Konzept zu erstellen und zu implementieren welches die Qualität der Korrekturmethode untereinander vergleicht und bewertet. Durch die Bewertung sollte es am Ende möglich sein Aussagen zu treffen wie: Die Korrekturmethode A ist besser als die Korrekturmethode B und C.

## 1 Einleitung

In Kärnten existieren circa 300 Wettermessstationen, die über das gesamte Bundesland verteilt sind (Stand Februar 2013). Bei einem Teil dieser Messstationen handelt es sich um hydrographische Messstationen, welche dazu dienen, die Pegelstände der für Kärnten relevanten Flüsse aufzuzeichnen und in weiterer Folge zu überwachen. Diese Messstationen sind besonders relevant für die Vorhersage von auftretendem Hochwasser. Anhand derer Messwerte muss im Falle einer Zunahme des Wasserstandes entschieden werden, ob etwa Schleusen geöffnet oder geschlossen bleiben.

Wie bei jeder Messung kann es auch beim Messen von Pegelständen zum Auftreten von Messfehlern kommen, welche in weitere Folge das Gesamtergebnis (Zeitreihe) beeinflussen bzw. verfälschen können. Es existiert eine Vielzahl von Faktoren die Messfehler verursachen können. Auftretender Frost, Verschleiß und mangelnde Stromversorgung sind nur einige wenige Ursachen die Messfehler erzeugen können. Um zu verhindern, dass fehlerhafte Messwerte das Gesamtergebnis verfälschen, bedarf es einer frühzeitigen Erkennung derer, um in weiterer Folge diese zu korrigieren. In Kärnten werden die vom Hydrographischen Dienst erfassten Messwerte innerhalb eines Zeitfensters von drei Minuten geprüft und bei Bedarf korrigiert. Da für jede Art von Fehler mehrere Korrekturmethode zu Verfügung stehen, stelle sich die Frage, welche dieser Korrekturmethode die besten Ergebnisse liefert. Um dies herauszufinden, wurde im Rahmen dieser Arbeit ein Java basiertes Tool

entwickelt welches die unterschiedlichen Korrekturmethode für hydrographische Zeitreihen untereinander vergleicht und auf objektive Weise bewertet.

## 2 Technischer Hintergrund

Im Vorfeld dieser Arbeit wurde im Rahmen einer Forschungskooperation zwischen dem Amt der Kärntner Landesregierung hydrographischer Dienst<sup>1</sup> und der Fachhochschule Kärnten ein gemeinsames Regelwerk (MOSER, 2010) erarbeitet, welches anwendbaren Methoden zur Messdatenprüfung bzw. zur Rekonstruktion von Daten dokumentiert. Das gesamte System, welches Messungen durchführt, diese speichert und in weiterer Folge vergleicht, besteht aus folgenden Komponenten. Wobei für unsere Betrachtung besonders die Softwares HyDaMS und ECO die wichtigsten Rollen spielen.

### 2.1 Hydrographisches Datenmanagementsystem (HyDaMS)

Bei HyDaMS handelt sich um eine in den 90er-Jahren von der Firma AquaPlan<sup>2</sup> entwickelten Software zur Analyse und Weiterverarbeitung von Zeitreihen. Mithilfe dieses Produkts ist es möglich, Stammdaten und Zeitreihen zu bearbeiten. HyDaMS besitzt eine einfache Benutzeroberfläche welche es dem Benutzer ermöglicht schnell auf Daten zuzugreifen. Außerdem ist diese Software so konzipiert, dass es kein Problem darstellt, diese modular zu erweitern. Des Weiteren ist es möglich, bereits vorhandene Zeitreihen zu manipulieren und zu bearbeiten.

### 2.2 ECO (Plug-In für HyDaMS)

Bei ECO handelt es sich um eine Software die Rahmen der Forschungskooperation zwischen der Fachhochschule Kärnten (Studienbereich Geoinformation & Umwelttechnologien) und dem Hydrographischen Dienst des Landes Kärnten. Die Aufgabe dieser Software ist es die von den Sensoren erfassten Daten zu prüfen und, sollten diese Messfehler aufweisen, in weiterer Folge zu korrigieren. Der Einsatz der jeweiligen Prüfprofile kann sowohl im manuellen als auch im zeitgesteuerten oder automatischen Betrieb erfolgen. Darüber hinaus ist es möglich, Prüfprofile selbst zu definieren und diese anzuwenden. Die Anwendung dieser Prüf- und Korrekturmethode kann für mehrere Zeitreihen erfolgen und bei Bedarf wiederholt erfolgen. Bei dem Einsatz von Prüfmethoden sind prinzipiell keine Grenzen gesetzt es kann sowohl nur ein Prüfprofil auf eine Zeitreihe angewandt werden, es können jedoch auch mehrere Prüfprofile auf eine Zeitreihe angewandt werden. ECO ist in HyDaMS integriert und besitzt eine einfache Benutzeroberfläche. Außerdem ist der Aufbau der Software so gegliedert, dass diese jederzeit erweitert werden kann. Es existieren drei verschiedene Korrekturmodi, die das ECO-System aufweist:

- *manuell* – die Korrektur erfolgt innerhalb eines zuvor definierten Zeitraums kann aber auch sofort eingeleitet werden;

---

<sup>1</sup> [http://www.ktn.gv.at/152237\\_DE-Organisation-Hydrographie](http://www.ktn.gv.at/152237_DE-Organisation-Hydrographie) (zuletzt besucht: 03.02.2013)

<sup>2</sup> <http://www.aquaplan.de/> (zuletzt besucht: 02.01.2013)

- *zeitgesteuert* – die Korrektur geschieht zu einem vordefinierten Zeitpunkt z. B. täglich 12:00 Uhr;
- *automatisiert* – die Korrektur findet statt wenn neue Messdaten (z. B. *Niederschlag*) eingehen und dies automatisiert vom System korrigiert werden.

### 2.3 Messstationen

In Kärnten gibt es circa 63 Messstationen, welche die Pegelstände von Flüssen messen. Diese Messstationen sind über das ganze Land Kärnten verteilt, eine kleine Anzahl an Messstationen befindet sich jedoch außerhalb des Landesgebietes. Die Gailitz etwa welche in der Nähe des Nevea-Sattels in Italien entspringt und ab in Arnoldstein in die Gail mündet beeinflusst die Hochwassersituation in Kärnten. Darum befinden sich schon am Oberarm der Gailitz Sensoren, welche die Durchflussrate messen, damit im Falle eines Pegelanstieges deren Daten schon rechtzeitig in die Hochwasserprognose des Landes einfließen. Weitere Messstationen befinden sich in Slowenien oder an der Grenze zu Italien. Einige Messstationen besitzen Sensoren für unterschiedliche Parameter manche hingegen messen nur einen einzigen Parameter. Mithilfe der Sensoren ist es möglich folgende Wetterphänomene zu messen: Temperatur, Niederschlag, Luftfeuchtigkeit, Luftdruck, Windgeschwindigkeit, Windrichtung, Abfluss, Schneehöhe und weitere. Jedoch kann nicht jede Messstation alle Parameter messen. Die Wartung und Kontrolle der Messstationen und deren Sensoren obliegen dem Land Kärnten.

### 2.4 Funktionsweise

Die Messstation misst entweder jeweils beim Eintreten des Wetterereignisses, z. B. Niederschlag oder kontinuierlich, wie dies beim Parameter Temperatur der Fall ist. Daraufhin werden die gemessenen Daten an die  $tstp^3$ -Schnittstelle gesandt danach werden diese mithilfe der Software ECO analysiert. Dabei wird erkannt ob es sich um fehlerhafte oder korrekte Daten handelt. Bei fehlerhaften Daten werden diese einer, bereits zuvor definierten, Korrektur unterzogen. Nachdem die Korrektur abgeschlossen ist, werden die korrigierten Daten an den  $tstp$ -Server gesandt und dort gespeichert. Die HyDaMS Software kann im Gegensatz zur ECO-Software direkt auf die Messdaten zugreifen. Jedoch dient die HyDaMS Software nicht zur Korrektur von Messdaten, sondern nur zur Betrachtung und Visualisierung der Messdaten.

## 3 Theoretischer Hintergrund

### 3.1 Fehlerarten

Die Fehlerarten, die auftreten können, sind äußerst vielfältig und teilweise Parameter abhängig. Allgemein kann gesagt werden, es gibt Fehler, die bei allen Parametern auftreten können, und parameterspezifische Fehler, die nur bei einigen Parametern auftreten. Des Weiteren können die Fehlerarten grob in drei Kategorien eingegrenzt werden:

---

<sup>3</sup> Time series transfer Protokoll

- *sensorbedingt* – Fehler treten aufgrund des Sensors auf z. B. Akkuleistung, Fehlkalibrierung eines Sensors;
- *wetterbedingt* – aufgrund von Wetterereignissen wie Sturmböen, Hagel, Frost treibenden Eisschollen kommt es zu Fehlern oder zur Beschädigung des Sensors;
- *umweltbedingt* – menschliche Aktivitäten um den Sensor führen zu Fehlern, z. B. vorbeifahrt eines Bootes → Pegelstand wird verfälscht.

Zurzeit (Februar 2013) existieren bereits 15 implementierte Fehlererkennungs- und deren Korrekturmethode für die Parameter Lufttemperatur, Niederschlag, Wasserstand, Abfluss, Schneehöhe, Windgeschwindigkeit und Windrichtung. Diese Fehlerkennungs- und Korrekturmethode erkennen und korrigieren unter anderem Ausreißer, Lücken, Flatliner und andere. In Zusammenarbeit mit dem Hydrographischen Dienst des Landes Kärnten wurde ein umfassendes Regelwerk (MOSER, 2010) erarbeitet, welches die häufigsten Fehler und deren passenden Korrekturmethode beinhaltet und deren genau Funktionsweise beschreibt.

### 3.2 Qualitätsbewertung

Als Kernpunkt dieser Arbeit gilt es, die Qualität der Korrekturmethode zu analysieren. Qualität ist jedoch ein äußerst dehnbarer Begriff, der unterschiedlich definiert werden kann. Um Qualität zu bewerten bzw. zu messen bedarf es zuvor einer passenden Definition, damit die Frage der Begrifflichkeit geklärt ist. Jene Korrekturmethode sind qualitativ am hochwertigsten einzustufen welche eine fehlerhafte Zeitreihe so korrigieren, dass diese der Originalzeitreihe am ähnlichsten ist. Somit gilt folgendes Optimum:

$$\text{Fehlerhafte-Zeitreihe} + \text{Korrekturmethode} = \text{Originalzeitreihe}$$

Um die korrigierte Zeitreihe mit der Originalzeitreihe zu vergleichen, wurden mehrere statistische Maßzahlen verwendet. Diese Maßzahlen vergleichen die korrigierte Zeitreihe mit der Originalzeitreihe und beschreiben deren Unterschiedlichkeit bzw. Ähnlichkeit auf mathematische Weise. Jede Maßzahl hat jedoch anderes Optimum definiert, um die Ähnlichkeit der korrigierten Zeitreihe mit der Originalen zu beschreiben. Da eine einzige Maßzahl nicht aussagekräftig genug ist, wurden mehrere Maßzahlen verwendet, um die Qualität noch besser zu beschreiben. Somit wird jene korrigierte Zeitreihe als am qualitativ hochwertigsten angesehen, welche die höchste Anzahl an Maßzahlen favorisiert.

#### Beispiel:

- Korrigierte Zeitreihe A (Korrekturmethode A);
  - 2 von 5 Maßzahlen favorisieren diese Korrekturmethode.
- Korrigierte Zeitreihe B (Korrekturmethode B);
  - 3 von 5 Maßzahlen favorisieren diese Korrekturmethode.

Somit ist die korrigierte Zeitreihe (B) als qualitativ hochwertiger anzusehen als korrigierte Zeitreihe (A).

### 3.2.1 Statistische Maßzahlen

#### *Bestimmtheitsmaß (ide)*

$$R^2 = \frac{[\sum_{i=1}^n (Q_i^* - \bar{Q}^*) \times (Q_i - \bar{Q})]^2}{\sum_{i=1}^n (Q_i - \bar{Q})^2 \times \sum_{i=1}^n (Q_i^* - \bar{Q}^*)^2} \quad (1)$$

Bei dem Bestimmtheitsmaß (1) handelt es sich um eine statistische Maßzahl zur Beschreibung der Variabilität zwischen beobachteten (Original) und modellierten (korrigierten) Werten (REED, ZHANG, & SCHAAKE, 2007). Das Ergebnis dieser Formel spiegelt den prozentuellen Unterschied zwischen Beobachteten und (GUPTA & SCHAEFLE, 2007) modellierten Werten wieder. Der Ergebnisbereich bewegt sich zwischen 0 und 1 wobei es sich bei 0 um das schlecht möglichste Ergebnis handelt und bei 1 um das bestmögliche Ergebnis (DRABEK, 2010).

#### *root mean squared error (rmse)*

$$rmse = \sqrt{\frac{1}{n} \sum_{i=1}^n (Q_i^* - Q_i)^2} \quad (2)$$

Bei dieser statistischen Maßzahl (2) handelt es sich um die Wurzel des mittleren quadratischen Fehlers *rmse*. GOURLEY & VIEUX (2006) beschreiben den gesamten Fehler sowohl zufällig als auch systematisch. Der Ergebniswert ist immer positiv umso höher dieser Wert umso schlechter das Ergebnis bzw. umso unähnlicher sind sich Original und korrigierte Zeitreihe. Ein Wert von 0 stellt das Optimum dar; Original und korrigierte Zeitreihe sind in diesem Fall mit hoher Wahrscheinlichkeit ident (DRABEK, 2010).

#### *normalised mean error (nme)*

$$nme = \frac{me}{\bar{Q}} = \frac{\frac{1}{n} \sum_{i=1}^n Q_i^* - Q_i}{\sum_{i=1}^n Q_i} \quad (3)$$

Der normierte mittlere Fehler (3) gibt ähnlich wie der *bias* (absoluter systematischer Fehler) Auskunft über den systematischen Fehler (GOURLEY & VIEUX, 2006). Durch die Normierung wird ein besser vergleichbarer Wert ermittelt. Das Ergebnis kann sowohl positiv als auch negativ sein. Das bestmögliche Ergebnis in Bezug auf deren Ähnlichkeit ist 0 (DRABEK, 2010).

#### *normalised standard deviation error (nsdve)*

$$nsdve = \frac{sdve}{\bar{Q}} = \frac{\sqrt{\frac{1}{n} \sum_{i=1}^n [(Q_i^* - Q_i) - me]^2}}{\bar{Q}} \quad (4)$$

Bei dieser Maßzahl (4) handelt es sich um die normierte Standardabweichung des Fehlers (REED, ZHANG, & SCHAAKE, 2007). Die Standardabweichung ist ein Maß für die Streuung

der Werte einer Zufallsvariable um ihren Mittelwert. Normiert mit dem Mittelwert der Beobachtung wird daraus eine besser vergleichbare Größe. Der  $nsdve$  ist immer positiv, für ein perfektes Modell ist  $nsdve=0$  (DRABEK, 2010).

#### ***Nash-Sutcliffe model efficiency (nsme)***

$$nsme = 1 - \frac{\sum_{i=1}^n (Q_i^* - Q_i)^2}{\sum_{i=1}^n (Q_i - \bar{Q})^2} \quad (5)$$

Das *Nash-Sutcliffe Kriterium* (5) kurz  $nsme$  ist ein kombiniertes Maß für zufälligen<sup>4</sup> und systematischen Fehler (NASH & SUTCLIFFE, 1970). Der Ergebniswert bewegt sich zwischen  $-\infty$  bis 1. Wobei 1 bedeutet, dass es sich hierbei um eine perfekte Modellierung handelt. Der Wert 0 bedeutet, dass die Modellierung so gut ist wie der Mittelwert der Beobachtung (Original).  $nsme < 0$  tritt hingegen auf, wenn der Mittelwert der Beobachtung ein besseres „Modell“, darstellt als das eigentliche Simulationsmodell (korrigierte Zeitreihe), d. h., wenn die Fehler des Modells (im Zähler) größer werden als die Varianz in der Beobachtung (im Nenner) (DRABEK, 2010).

#### ***Benchmark-Modell (bme)***

$$bme = \frac{\sum_{i=1}^n (Q_i^* - Q_i)^2}{\sum_{i=1}^n (Q_i - Q_b)^2} \quad (6)$$

Das Benchmark-Modell (6) ist eine Erweiterung des Nash-Sutcliffe Kriteriums und unterscheidet sich im Wesentlichen nur durch eine andere Definition der Variablen  $\bar{Q}$ .

#### **Nash-Sutcliffe versus Benchmark-Modell**

Das Nash-Sutcliffe Kriterium benötigt zu dessen Berechnung immer den Mittelwert der Beobachtung  $\bar{Q}$ . Dieser ist jedoch bei der Berechnung von Wetterdaten bzw. deren Vergleich nicht unbedingt genau. Man berechne zum Beispiel den Jahresmittelwert des Pegelstands eines Flusses für ein ganzes Jahr. Diesen Wert als Mittelwert für die Berechnungen etwa des Monats Augusts oder Dezembers heranzunehmen, ist unbrauchbar. Somit erwähnten schon GUPTA & SCHAEFLE (2007), dass es sinnvoller ist, nicht etwa den Jahresmittelwert  $\bar{Q}$  heranzuziehen, sondern dem Mittelwert des jeweils zu berechnenden Monats  $Q_b$ . In unserem Fall jeweils den Mittelwert der Monate August oder Dezember. Deshalb unterschieden sich die zwei Formeln nur durch eine Variable:  $\bar{Q}$  durch  $Q_b$ .

<sup>4</sup> *Systematischer und zufälliger Fehler* [http://www.statistics4u.info/fundstat\\_germ/ee\\_precision\\_accuracy.html](http://www.statistics4u.info/fundstat_germ/ee_precision_accuracy.html) (zuletzt besucht:01.02.2013)

### 3.2.2 Optimum und Pessimum der Maßzahlen

Wie in der Tabelle 1 erkennbar reichen die Grenzwerte bei der Maßzahl nme von  $-\infty$  bis  $+\infty$  wobei 0 das Optimum darstellt. Umso weiter das Ergebnis von 0 entfernt ist umso schlechter das Ergebnis. Die Grenzwerte der Maßzahlen rsme, nsdve befinden sich zwischen 0 und  $+\infty$ , wobei 0 das Optimum darstellt. Umso höher der Wert umso schlechter das erzielte Ergebnis. Bei der Maßzahl nsme ergibt sich wieder eine andere Konstellation der Wert 0 bedeutet, dass das erzielte Ergebnis gleich gut ist wie der Mittelwert der Beobachtung. Das Optimum stellt der Wert 1 dar.

**Tabelle 1:** Optimum und Pessimum der Maßzahlen

Maßzahlen	Optimum	Pessimum
root mean squared error (rsme)	0	höher
Nash-Sutcliffe model efficiency (nsme)	1	$x=0$ Ergebnis gleich gut wie Mittelwert
Bestimmtheitsmaß(ide)	1	0
normalised mean error (nme)	0	niedriger oder höher als 0
normalised standard deviation error (nsdve)	0	höher

## 4 Implementierung

Zur Ermittlung der Qualität der einzelnen Korrekturmethode wurde ein Programm in Java implementiert welches die unterschiedlichen bereits korrigierten Zeitreihen mit der original Zeitreihe vergleicht.

Der Qualitätstest erfolgt in mehreren Schritten:

1. Auswählen einer Zeitreihe z. B. *Drau-Amlach-Pegelstand*.
2. Mithilfe der HyDaMS-Software wird die Zeitreihe betrachtet und Fehler werden erzeugt. Es wird z. B. ein Teil der Zeitreihe gelöscht.
3. Die ECO-Software wird gestartet und eine Korrekturmethode für dies Zeitreihe wird ausgewählt, z. B. *limitcheck.xml*.
4. Die Zeitreihe *Drau-Amlach-Pegelstand* wird geprüft und korrigiert.
5. Es werden weitere Korrekturmethode auf diese Zeitreihe angewandt.
6. Das Quality-Tool evaluiert daraufhin, welche der Korrekturmethode die Zeitreihe am besten korrigiert hat.
7. Analyse der Ergebnisse

Anhand der durch das Programm (*Quality Tool*) berechneten Werte ist nun durch den Benutzer zu erkennen, welche Korrekturmethode die Beste ist. Die Interpretation der Ergebnisse erfolgt durch den Benutzer und nicht durch das Programm. Die genaue Funktionswei-

se des Programms und der Aufbau der programmierten Klassen ist dem technischen Bericht zu entnehmen (ROBATSCH & LEHRER 2013)

## 5 Anwendung

Im nächsten Schritt wurde das Implementierte *Quality Tool* getestet, um den Einsatz der einzelnen statistischen Maßzahlen möglichst objektiv zu bewerten wurden Testdaten verwendet. Die Ergebnisse wurden mithilfe des Programms *Microsoft Excel* grafisch aufbereitet, um einfacher interpretiert werden zu können.

**Tabelle 2:** Testdaten grafisch aufbereitet (Originaldaten Spalte 1, Messzeitpunkt Spalte 2, Spalte 4 manipulierte Originaldaten, Spalte 5-8 korrigierte Zeitreihen)

Töbriug								
Temperatur in °C								
Zeitraum								
VON 0:00 12.09.2012 BIS 03:00 12.09.2012								
Originaldaten	Messzeitpunkt	Art (Original)	Original manipuliert	Korrigiert A	Korrigiert B	Korrigiert C	Korrigiert D	Art (Korrigiert)
5	○ 00:00	Korrekt	5	5	5	5	5	5 Original
7	○ 00:15	Korrekt	7	7	7	7	7	7 Original
7	○ 00:30	Korrekt	7	7	7	7	7	7 Original
6	○ 00:45	Korrekt	6	6	6	6	6	6 Original
8	○ 01:00	Korrekt	8	8	8	8	8	8 Original
10	○ 01:15	Korrekt	10	10	10	10	10	10 Original
11	○ 01:30	synthetischer Fehler	0	10	9	11	9	9 Korrigiert
10	○ 01:45	synthetischer Fehler	0	10	9	12	10	10 Korrigiert
11	○ 02:00	synthetischer Fehler	0	9	8	10	9	9 Korrigiert
12	○ 02:15	Korrekt	12	12	12	12	12	12 Original
10	○ 02:30	Korrekt	10	10	10	10	10	10 Original
7	○ 02:45	Korrekt	7	7	7	7	7	7 Original
6	● 03:00	Korrekt	6	6	6	6	6	6 Original

Wie in Tabelle 2 erkennbar ist, befinden sich in der ersten Spalte von links (Bezeichnung: Originaldaten) jeweils die Ergebnisse der Originalzeitreihe. In der zweiten Spalte (Bezeichnung: Messzeitpunkt) werden die einzelnen Messzeitpunkte der Messungen angegeben. Die dritte Spalte (Bezeichnung: Art (Original)) gibt Auskunft über die Art der jeweiligen Spalten. Korrekt: Bedeutet, dass es sich hierbei um fehlerfreie Messwerte handelt. Synthetischer Fehler: Bedeutet die Werte, welche sich in den Spalten rechts (Bezeichnung: Original manipuliert) dieser Spalte und Zeile befinden, mithilfe von HyDaMS manipuliert wurden. Wie zu erkennen sind diese Werte je mit dem Wert 0 °C belegt. Die Spalten fünf bis acht stellen die Ergebnisse der einzelnen Korrekturmethode, A-B-C-D, dar. Die Korrekturmethode wurden bewusst nur mit A, B, C, D bezeichnet um die Objektivität der Analyse zu erhöhen. Denn so ist es für den Bewerter der Analyse schwieriger, etwa eine favorisierte Korrekturmethode zu bevorzugen. Die letzte Spalte wiederum kennzeichnet (Bezeichnung: Art (Korrigiert)) ob es sich hier um Originalteil oder den von der Korrekturmethode korrigierten Teil handelt. „Original“: Steht für den nicht korrigierten Teil; „Korrigiert“: Steht für den korrigierten Teil der Zeitreihe.



**Tabelle 3:** Ergebnisse der Analyse der Testdaten (die besten Werte sind grau markiert)

Maßzahl	Original Manipuliert	A	B	C	D
rsme	5129102485590660,00	0.6201736729460423	0.8257382111239755	0.6201736729460423	0.7844645405527362
nsme	-4408759124087590,00	0.9209245742092458	0.7785888077858881	0.9209245742092458	0.8734793187347932
ide	-687612398180,47	0.016815156719622155	0.016846503755421557	0.01456745066234117	0.01689227362093512
rme	-0.2909090909090909	-0.02727272727272727	-0.05454545454545454	0.00909090909090909	-0.03636363636363636
nsdve	5309993696592420,00	435943926942464,00	4457772486028200,00	4228328313953220,00	4392217008292490,00

In Tabelle 3 werden die Ergebnisse der Analysen, durch die statistischen Maßzahlen, dargestellt. In der ersten Spalte (Bezeichnung: Maßzahl) werden die angewendeten Maßzahlen aufgelistet. In der zweiten Spalte (Bezeichnung: Original Manipuliert) sind die Ergebnisse zu erkennen welche die manipulierte Zeitreihe verglichen mit der Originalzeitreihe erreicht. Hierbei ist zu erkennen, dass die manipulierte Zeitreihe bei jeder statistischen Maßzahl am schlechtesten abschneidet. Spalten drei bis sechs stellt die Ergebnisse der restlichen Korrekturmethode verglichen mit der Originalzeitreihe dar. Die beste(n) Korrekturmethode wurde jeweils pro Zeile *grau markiert*. Bei der schlussendlichen Bewertung der Ergebnisse ist zu erkennen, dass sich die Korrekturmethode A und C ein Kopf an Kopf liefern. Erst ab der vierten Maßzahl entscheidet die Korrekturmethode C das Rennen für sich. Die Korrekturmethode *ide* (Bestimmtheitsmaß) wurde nicht in die Bewertung mit einbezogen, siehe Fazit.

## 6 Diskussion

Zu Beginn der Arbeit wurde die Idee angedacht, die einzelnen Maßzahlen zu einer Gesamtformel zusammenzufügen. Diese Idee wurde jedoch während der Projektarbeit wieder verworfen, da die Maßzahlen untereinander zu unterschiedliche Grenzwerte aufweisen. In der Folge der weiteren Ausbildung wurde jedoch eine Möglichkeit gefunden, diese Maßzahlen zu einer Zahl zu vereinigen. Um dies anzustellen, ist es jedoch notwendig, die unterschiedlichen Optima und Pessima zu vereinheitlichen. Dies kann mithilfe der Definition von Cost- und Benefit-Criteria erfolgen. Das Programm DEFINITE<sup>5</sup> könnte sich bei solch einer Vereinheitlichung der Maßzahlen als besonders hilfreich erweisen. Eine solch vereinheitlichte Zahl könnte vom Programm selbst interpretiert und ausgewertet werden. Ferner ist es noch notwendig das Programm noch zusätzlichen Tests mit unterschiedlichsten Testdaten zu unterziehen, um dessen Aussagekraft weiter zu untermauern. Zusätzlich könnte in Zukunft auch noch das Interface des Programms verbessert und benutzer-freundlicher gestaltet werden.

<sup>5</sup> <http://www.ivm.vu.nl/en/projects/Projects/spatial-analysis/DEFINITE/index.asp> (zuletzt besucht: 19.01.2014)

## 7 Fazit

Zur Bewertung der Qualität der Korrekturmethode mithilfe der ausgewählten Maßzahlen ist das entwickelte Programm durchaus zweckdienlich. Die Maßzahl des Bestimmtheitsmaßes ist aufgrund der Ergebnisse am wenigsten geeignet. Zwar erzielt diese Maßzahl auch gute Ergebnisse, jedoch sollte man nicht außer Acht lassen, dass das Bestimmtheitsmaß nur ein deskriptives Maß für die Güte der Anpassung der Regressionsgeraden ist und kein Maß für die Qualität der Schätzung.

## Literatur

- DRABEK, U. (2010), Anwendungsbezogene Aspekte der operationellen Durchflussvorhersage (Bd. 221). (I. f. o. Univ.-Prof. Dipl.-Ing. Dr. Günter Blöschl, Hrsg.) Wien, Österreich, TU Wien.
- GOURLEY & VIEUX (2006), A method for identifying sources of model uncertainty in rainfall-runoff simulations. *Journal of Hydrology*, 327, 68-80.
- GUPTA & SCHAEFLI (2007), *Hydrological Processes*. John Wiley & Sons, Ltd., USA:
- MOSER, J. E. (2010), Theoretisches Regelwerk – Fehlererkennungs- und Rekonstruktionsmethoden für meteorologische/hydrologische Messungen. Klagenfurt: Amt der Kärntner Landesregierung/Abteilung 8/Hydrographischer Dienst.
- NASH & SUTCLIFFE (1970), River flow forecasting through conceptual models part I – A discussion of principles. *Journal of Hydrology*, 10 (3), 282-290.
- REED, S., ZHANG, Z. & SCHAAKE, J. (2007), A distributed hydrological model and threshold frequency-based method for flash flood forecasting at ungauged stations. *Journal of Hydrology*, 337, 402-420.
- ROBATSCH, E. (2013), ECO-Quality of Service. Bachelorarbeit 1.
- ROBATSCH & LEHRER (2013), Technischer Bericht – ECO-Quality of Service.