


Schriftenreihe CIplus, Band 3/2013

Herausgeber: T. Bartz-Beielstein, W. Konen, H. Stenzel, B. Naujoks



FIWA - Methoden der Computational Intelligence für Vorhersagemodelle in der Finanz- und Wasserwirtschaft (Schlussbericht)

Thomas Bartz-Beielstein und Oliver Flasch



Fachhochschule Köln
Cologne University of Applied Sciences

FIWA

Methoden der Computational Intelligence für Vorhersagemodelle in der Finanz- und Wasserwirtschaft

(Förderkennzeichen 17N2309)

Schlussbericht

Förderlinie „Ingenieur*Nachwuchs*“ 2009 (Informatik)
im Rahmen des Programms „Forschung an Fachhochschulen“

Prof. Dr. Thomas Bartz-Beielstein

Institut für Informatik

Fakultät für Informatik und Ingenieurwissenschaften

Fachhochschule Köln

Inhaltsverzeichnis

1	Allgemeines	2
2	Vorhaben	3
2.1	Aufgabenstellung	3
2.2	Voraussetzungen	4
2.3	Planung und Ablauf	5
2.4	Wissenschaftlich-technischer Stand	7
2.5	Zusammenarbeit mit anderen Stellen	11
3	Wissenschaftlich-technische Ergebnisse	11
3.1	Modul M-1 „Genetic Programming“	12
3.1.1	GP-1: Verallgemeinerung und Dokumentation des GP-Systems	12
3.1.2	GP-2: Entwicklung weiterer allgemeiner Zeitreihenanalysefunktionen	13
3.2	Modul M-2 „Finanzwirtschaft“	16
3.2.1	F-1: Entwicklung von auf Finanzzeitreihen spezialisierten Analysefunktionen	16
3.2.2	F-2: Optimierung der Prognosemodellauswertung	16
3.2.3	F-3: Entwicklung eines Bewertungswerkzeugs für Finanzzeitreihen-Prognosemodelle	16
3.2.4	F-4: Evaluation der Prognosegenauigkeit für Finanzzeitreihen	18
3.2.5	CS-1: Vergleichende Übersicht: Prognosemodelle in der Finanzwirtschaft	18
3.2.6	CS-2: Vergleich des FIWA-GP Prognosemodells mit anderen Modellen in der Finanzwirtschaft	18
3.2.7	MT-1: Erstellung von Softwarepaketen für Finanzzeitreihen	22
3.3	Modul M-3 „Wasserwirtschaft“	22
3.3.1	W-1: Datenerfassung und GP-Anpassung für Daten aus der Wasserwirtschaft	22

3.3.2	W-2: Entwicklung von auf Wasserwirtschaftszeitreihen angepassten Analysefunktionen	23
3.3.3	W-3: Entwicklung eines Bewertungswerkzeugs für Wasserwirtschafts-Prognosemodelle	24
3.3.4	W-4: Evaluation der erreichten Prognosegenauigkeit für Wasserwirtschaftszeitreihen	24
3.3.5	CS-3: Vergleichende Übersicht: Prognosemodelle in der Wasserwirtschaft . . .	26
3.3.6	CS-4: Vergleich des FIWA-GP Prognosemodells mit anderen Modellen in der Wasserwirtschaft	26
3.3.7	MT-2: Erstellung von Softwarepaketen für Wasserdaten	29
3.4	Modul M-4 „Generalisierbarkeit“	29
3.4.1	G-1: Unterschiede und Gemeinsamkeiten von GP-Systemen in der Finanz- und Wasserwirtschaft	29
3.4.2	G-2: Generalisierbarkeit	30
3.5	Spezielle Arbeitspakete	30
3.6	Wissenschaftliche Publikationen	31
4	Voraussichtlicher Nutzen	31
5	Ergebnisse Dritter	32
6	Veröffentlichungen im Projekt	34
7	Literatur	39
	Abkürzungsverzeichnis	42

1 | Allgemeines

Dieser Schlussbericht beschreibt die im Projekt „Methoden der Computational Intelligence für Vorhersagemodelle in der Finanz- und Wasserwirtschaft“ (FIWA) im Zeitraum von Juni 2009 bis einschließlich November 2012 erzielten Ergebnisse.

Zuwendungsempfänger	Institut für Informatik, Fakultät Informatik und Ingenieurwissenschaften, Fachhochschule Köln
Projektleiter	Prof. Dr. Thomas Bartz-Beielstein
Förderkennzeichen	17N2309
Vorhabensbezeichnung	Methoden der Computational Intelligence für Vorhersagemodelle in der Finanz- und Wasserwirtschaft (FIWA)
Laufzeit des Vorhabens	06/2009 – 11/2012
Berichtszeitraum	06/2009 – 11/2012
Projektpartner	Dortmund Intelligence Project (DIP) GmbH, Technische Werke Emmerich am Rhein GmbH (TWE), TU Dortmund, Quaesta Capital GmbH, Vrije Universiteit Amsterdam

Der Bericht ist wie folgt gegliedert: Abschnitt 2 gibt einen Überblick über Aufgabenstellung, Voraussetzungen, Planung und Verlauf, wissenschaftlich-technischen Stand und über die Zusammenarbeit mit Dritten im Projekt FIWA. Abschnitt 3 beschreibt die erzielten wissenschaftlich-technischen Ergebnisse. Abschnitt 4 vergleicht den aktuellen Stand des Vorhabens mit der ursprünglichen Arbeits-, Zeit- und Ausgabenplanung. Abschnitt 5 beschreibt Informationsrecherchen nach projektrelevanten Ergebnissen Dritter. Abschnitt 6 führt alle im Projekt erfolgten Veröffentlichungen auf.

2 | Vorhaben

2.1 Aufgabenstellung

Das Projekt FIWA befasste sich mit der Entwicklung modularer Systeme zur Analyse und Prognose von Daten aus der Finanz- und Wasserwirtschaft mittels Verfahren der Computational Intelligence (CI). Dabei liegt der methodische Fokus auf Ansätzen aus dem CI-Unterbereich Genetic Programming (GP).

Sowohl in der Wasserwirtschaft als auch in der Finanzwirtschaft sind genauere Prognosen von Zeitreihen notwendig. Zukünftige Entwicklungen aus bestehenden Trends zu bestimmen kann Entscheidungs-, Handlungs- und Kontrollstrategien mit dem notwendigen Wissen versorgen um erfolgreich zu sein. Sowohl in industriellen Prozessen als auch in ökonomischen Prozessen tauchen dabei häufig Probleme auf, deren Lösung die Berücksichtigung der hohen Komplexität und nichtlinearen Dynamik benötigen. Besonders für längerfristige Vorhersagen ergibt sich eine inhärent wachsende Unsicherheit. Rasch veränderliche Einflüsse aus Umgebung, Umwelt, Politik und Ökonomie sind zusätzlich zu beachten.

Auf Grund dieser Schwierigkeiten sind klassische Methoden im Bereich der Zeitreihenprognose nur schlecht anzuwenden. Ziel dieses Vorhabens ist deshalb die Entwicklung und Anwendung moderner CI Methoden, vor allem im Bereich GP, die potentiell selbst komplexe Zusammenhänge in Zeitreihen gut abbilden können. Mit GP lassen sich zudem auch Komponenten der klassischen Zeitreihenprognose in ein neues Modellierungsframework integrieren. Damit ergibt sich die Möglichkeit automatisiert sowohl klassisch lösbare als auch hochkomplexe Zeitreihen zu modellieren.

Im Projekt FIWA sollen die entwickelten GP-Systeme folglich für die Modellierung von Finanzzeitreihen (z.B. Aktienkurse) und Zeitreihen aus der Wasserwirtschaft (z.B. Füllstände von Reservoirs) verwendet werden. Dabei soll auch die Möglichkeit untersucht werden, neben den zu modellierenden Größen auch weitere Eingabedaten zu berücksichtigen (z.B. Text von Finanznachrichten oder Wetterdaten).

Konkret werden aus dem Bereich der Finanzwirtschaft Daten der beteiligten KMU (Dortmund Intelligence Project (DIP) GmbH und Quaesta Capital GmbH) zur Verfügung gestellt. Auf bereits bestehenden Erfahrungen in GP und klassischen Methoden soll aufgebaut werden. Mit der Verbesserung dieser Methodik soll es den Projektpartnern ermöglicht werden, optimierte GP-Systeme zur Portfoliooptimierung oder zum Management von Accounts, Fonds oder strukturierten Produkten für alternative Investments in der Finanzwirtschaft einzusetzen.

Im Bereich der Wasserwirtschaft ist der Einsatz besonders innovativ, da hier der Einsatz von GP

erstmalig vorgenommen wurde. In diesem Bereich sind zugrunde liegende Prozessmodelle selten bekannt. Daher wird im Projekt FIWA auf datengetriebene Modellierung mit GP zurückgegriffen. Der Technische Werke Emmerich am Rhein GmbH (TWE) soll damit die Möglichkeit geboten werden vorausschauende Steuer- und Regelungsaufgaben in abwassertechnischen Anlagen zu realisieren.

Durch Veröffentlichungen und Kooperationen mit führenden Wissenschaftlern und Experten (Prof. Rudolph (TU Dortmund), Prof. Kleijnen (Tilburg University) und Prof. Eiben (Vrije Universiteit Amsterdam), Dr. Vladislavleva (Evolved Analytics)) aus den Bereichen CI und GP wird des Weiteren sichergestellt, dass auch der akademische Nutzen der im Projekt FIWA durchgeführten Forschung erzielt wird.

2.2 Voraussetzungen

Das Projekt FIWA hatte eine Gesamtlaufzeit von 3 Jahren und 6 Monaten (01.06.2009 bis zum 30.11.2012) und wurde mit einer Gesamtsumme von 258.424,00 € (exklusive 10% Projektpauschale) gefördert. Davon entfielen

- 224.506,00 € auf Personalausgaben,
- 22.456,00 € auf sächliche Verwaltungsausgaben und
- 11.462,00 € auf Gegenstände über 410 Euro (Investitionen).

Inklusive der Projektpauschale von 10% betrug die Fördersumme also 270.737,60 €.

Durchgeführt wurde das Projekt am Institut für Informatik an der Fakultät für Informatik und Ingenieurwissenschaften am Campus Gummersbach der Fachhochschule Köln. An dieser Fakultät hat der Projektleiter seit Oktober 2006 eine Professur für Angewandte Mathematik inne. Zu Projektbeginn hatte er mehr als 50 Publikationen im Bereich Computational Intelligence (CI) veröffentlicht. Er arbeitete von 2000–2006 im Sonderforschungsbereich „Design und Management komplexer technischer Prozesse und Systeme mit Methoden der Computational Intelligence“ an der *TU Dortmund* in einem Projekt zur Umsetzung der Grundlagenresultate in praxisorientierte Algorithmen und Werkzeuge. Die vom Projektleiter entwickelte Sequentielle Parameteroptimierung (SPO) wurde als wichtiges Werkzeug zur Entwicklung einer Variante von GP, der algorithmischen Chemie [83, 82], erfolgreich eingesetzt¹. Die Zusammenhänge wurden in Bartz-Beielstein u. a. [64] veröffentlicht. Diese Zusammenarbeit führte zu weiteren Publikationen, in denen das Zusammenspiel von GP und SPO und von SPO–Anwendungen untersucht wurden [81, 87, 65, 86].

Der Projektleiter ist Mitglied in den Programmkomitees der führenden Konferenzen und Zeitschriften auf dem Gebiet der CI. Bereits zu Projektbeginn existierte eine intensive Zusammenarbeit des Projektleiters mit Forschern auf internationaler Ebene, unter anderem mit den Universitäten Antwerpen, Tilburg, Eindhoven und Leeds. Weitere Forschungsthemen, die in dieser Arbeitsgruppe behandelt werden, stammen aus den Bereichen DoE, Simulation und Optimierung.

Ab Projektbeginn stand Herr Dipl.-Inform. Oliver Flasch dem Projekt als verantwortlicher Sachbearbeiter in Vollzeit zur Verfügung. Herr Flasch besitzt ein Diplom der TU Dortmund und war vor Projektbeginn bei der *DIP GmbH* angestellt. Er besitzt hervorragende Kenntnisse im Bereich CI und war an der Entwicklung kommerzieller GP-Systeme beteiligt. Seine Zulassung zur Promotion des

¹SPO ist in Form des open-source Pakets Sequential Parameter Optimization Toolbox (SPOT) frei verfügbar. Siehe auch www.cran.r-project.org/package=SPOT

Arbeitstitels „Methoden der Computational Intelligence für Vorhersagemodelle in der Finanz- und Wasserwirtschaft“ in Kooperation mit der *TU Dortmund* war zu Projektbeginn bereits genehmigt.

Die am Projekt beteiligten *KMU DIP GmbH* und *Quaesta Capital GmbH* hatten bereits sehr gute Erfahrungen mit dem Einsatz von GP für die Zeitreihenprognose. Deren GP-Systeme gehörten weltweit zu den Ersten im praktischen Einsatz in der Finanzwirtschaft. Zu Projektbeginn befanden sich keine GP-Systeme im praktischen Einsatz in der Wasserwirtschaft. Durch eine mehrjährige Kooperation mit der *DIP GmbH* und *Quaesta Capital GmbH* hat der Projektleiter gute Kontakte im Bereich der deutschen Finanzdienstleister. Gleichzeitig konnte das kommerzielle GP-System *vTrader* der *DIP GmbH* vor Projektbeginn an die Bedürfnisse der Branche angepasst und der erfolgreiche Markteintritt erzielt werden. Die *DIP GmbH* und *Quaesta Capital GmbH* stellte dem Projekt seit Beginn bei Bedarf tagesaktuelle Finanzzeitreihen kostenlos zur Verfügung und stand dem Projekt mit ihrem umfassenden Erfahrungsschatz in der Finanzwirtschaft beratend zur Seite. Im Bereich Wasserwirtschaft erklärte sich die *Technische Werke Emmerich am Rhein GmbH* dazu bereit, dem Projekt Prozessdaten aus Kläranlagen über einen Zeitraum von mehreren Monaten für eine Analyse und Erprobung zur Verfügung zu stellen.

Prof. Günter Rudolph (Lehrstuhl für Algorithm Engineering, *TU Dortmund*) wendet CI-Verfahren zur Mehrzieloptimierung und industriellen Optimierung an. Prof. Rudolph betreut das Promotionsvorhaben des Projekts an der *TU Dortmund*. Ein langjähriger Kontakt des Projektleiters zu Prof. Agoston E. Eiben (Head of the Computational Intelligence Group, Head of the Department of Computer Science an der *Vrije Universiteit Amsterdam*) besteht im Bereich CI und der experimentellen Analyse von Algorithmen. Dieser Kontakt wurde seit Projektbeginn zum inhaltlichen Austausch genutzt. Durch diesen Kontakt besteht zusätzlich die Möglichkeit für Absolventen der FH Köln, eine Promotion an der *Vrije Universiteit Amsterdam* durchzuführen. Über die Kompetenzplattform *Sustainable Technologies and Computational Services for Environmental and Production Processes (STEPS)* bestand an der Fachhochschule Köln schon vor Projektbeginn ein Kolleg für alle Promovenden der Hochschule.

2.3 Planung und Ablauf

Das Projekt wurde in vier Module mit insgesamt zwölf Arbeitspaketen aufgeteilt. Die Dimensionierung der Arbeitspakete wurde so gewählt, dass sie innerhalb eines Semesters zu bearbeiten waren (siehe Abbildung 1). Am Ende jeden Semesters wurde intern kontrolliert, ob die definierten Ziele erreicht wurden. Neben diesen projektbegleitenden Kontrollen wurden die folgenden Meilensteine definiert:

MS-1: Abnahme des erweiterten GP-Systems für Wasserdaten. – Q4 / 2009 Mit dem planmäßigen Erreichen des Meilensteins MS-1 wurde erweiterte Funktionalität korrekt in das GP-System integriert, so dass auch Daten aus der Wasserwirtschaft modelliert werden konnten. Durch die Portierung des Kerns des *vTrader*-Systems in die beliebte und weit verbreitete freie Statistik-Softwareumgebung R konnte der Implementierungsaufwand durch die Nutzung von Open Source Software verringert werden. Gleichzeitig steht mit RGP nun erstmals ein leistungsfähiges freies GP-System in R zur Verfügung.²

²RGP steht unter <http://rsymbolic.org> zum freien Download bereit und wird auch nach Projektende aktiv weiterentwickelt.

MS-2: Abnahme Analysefunktionen für die Wasserwirtschaft – Q2 / 2010 Mit dem planmäßigen Erreichen des Meilensteins MS-2 wurde sichergestellt, dass die Schnittstellen für Daten aus der Wasserwirtschaft korrekt spezifiziert wurden, so dass im zweiten Teil des Arbeitspakets W-2 die Implementierung und die Tests durchgeführt werden konnten. Daten aus der Wasserwirtschaft konnten korrekt eingelesen werden und stehen nach einer Datenvorverarbeitung in einem standardisierten Format zur Verfügung. Dabei handelt es sich zum Beispiel um Messwerte wie Niederschlagsmengen und Füllstände in Regenüberlaufbecken.

MS-3: Abnahme des erweiterten GP-Systems – Q3 / 2010 Durch das planmäßige Erreichen des Meilensteins MS-3 wurden die Daten und Analysefunktionen für die Wasserwirtschaft (W-2) und die zusätzlichen Funktionen für das GP-System (GP-2) korrekt implementiert, so dass eine Freigabe der Funktionen erfolgen konnte. Das für eine Anwendung in der Wasserwirtschaft spezialisierte INT2-Modell wurde in RGP integriert und Ergebnisse dieser Integration wurden auf wissenschaftlichen Konferenzen veröffentlicht [30].

MS-4: Festlegung der Prognosemodelle – Q1 / 2011 Dieser Meilenstein diente der Auswahl der Modelle und der Abnahme der Ergebnisse aus den Case-Studies CS-1 und CS-3. Der große Umfang dieser Case-Studies bedingte eine Abweichung von der ursprünglichen Planung. CS-1 („Vergleichende Übersicht: Prognosemodelle in der Finanzwirtschaft“) wurde als Masterarbeit MT-1 („Algorithmic Stock Trading as a Real-World Test Problem for Time Series Prediction Methods“) vergeben und Anfang 2011 erfolgreich abgeschlossen. CS-3 („Vergleichende Übersicht: Prognosemodelle in der Wasserwirtschaft“) wurde bereits in 2009 im Masterstudiengang „Automation & IT“ von drei Studenten durchgeführt und erfolgreich abgeschlossen.

MS-5: Bewertung der Vorhersagegüte für Daten der Finanz- und Wasserwirtschaft. – Q3 / 2011 Mit dem Erreichen dieses Meilensteins wurde sichergestellt, dass die in den Paketen F-4 und W-4 entwickelten Prognose- und Bewertungswerkzeuge korrekt entwickelt wurden. Durch die Ausrichtung der GECCO INDUSTRIAL CHALLENGE³ 2011, eines internationalen Wettbewerbs für Prognosesysteme für die Finanzwirtschaft, konnte die Evaluation der Prognosewerkzeuge in erweitertem Umfang erfolgen. So erfolgte nicht nur ein Vergleich mit klassischen Methoden der Zeitreihenprognose, sondern auch mit modernsten Methoden aus der CI.

MS-6: Abnahme der Ergebnisse aus den Vergleichen – Q1 / 2012 Mit dem Erreichen dieses Meilensteins wurde überprüft, ob die Vergleiche in den Case-Studies CS-2 und CS-4 korrekt durchgeführt wurden. Für die Bearbeitung der Case-Study CS-4 („Vergleich des FIWA-GP Prognosemodells mit anderen Modellen in der Wasserwirtschaft“) sowie für die Masterarbeit MT-2 („Erstellung von Softwarepaketen für Wasserdaten“) konnten keine geeigneten Studierenden gewonnen werden. Daher mussten die für das Erreichen der Meilensteine notwendigen Arbeiten von dem im Projekt beschäftigten wissenschaftlichen Mitarbeiter übernommen werden. Durch den gesteigerten Arbeitsaufwand

³Bei der Genetic and Evolutionary Computation Conference (GECCO) handelt es sich um eine der wichtigsten Konferenzen des Forschungsbereiches.

ergab sich eine Verzögerung beim Erreichen des Meilensteins.

Im letzten Semester des Projekts kam es zu einem äußerst starken und im Vorfeld nicht kalkulierbaren Anstieg der Studierendenzahlen. Mit 863 Erstsemestern war die Zahl im Winterhalbjahr gegenüber 2010 um fast die Hälfte angestiegen. Durch den erhöhten Betreuungsaufwand in den Lehrveranstaltungen kam es zu Verzögerungen bei der Bearbeitung der Arbeitspakete G-1 und G-2. Aus diesen Gründen wurde eine kostenneutrale Verlängerung des Projekts bis zum 30.11.2012 beantragt und genehmigt, was die Erreichung des Meilensteins MS-6 und damit einen erfolgreichen Projektabschluss ermöglichte. Die ursprüngliche Zeit- und Meilensteinplanung kann der Vorhabensbeschreibung des Projekts entnommen werden.

Abbildung 1 stellt den Projektverlauf in Form eines Gantt-Charts dar. Für die Arbeitspakete wurden die folgenden Bezeichnungen verwendet: **GP**: Genetic Programming, **F**: Finanzwirtschaft, **CS**: Case-Studies, **MT**: Master-Theses, **W**: Wasserwirtschaft, **G**: Generalisierbarkeit, **D**: Dissertation und **K**: Koordinierung. **MS** bezeichnet einen Meilenstein.

2.4 Wissenschaftlich-technischer Stand

Die im Projekt FIWA behandelten Aufgabenstellungen lassen sich mit empirischen Modellierungsverfahren bearbeiten. In der Praxis werden Verfahren der linearen und nichtlinearen Regression, NN, SVM und viele weitere Verfahren sowie Kombinationen dieser Verfahren eingesetzt. Der Schwerpunkt des Antrags FIWA lag im Bereich der symbolischen Regression mittels GP, so dass sich die Beschreibung an dieser Stelle auf dieses Gebiet konzentriert. Schlittgen [93], Brockwell u. Davis [67] und Tsay [97] stellen den aktuellen Stand der Forschung im Bereich Zeitreihenanalyse und -prognose dar.

Genetic Programming Poli u. a. [89] präsentiert einen Überblick und eine umfassende Darstellung der GP-Forschung. Weitere GP-Operatoren bzw. Variationen bestehender Operatoren werden ständig weiter entwickelt. Lineares GP, Graph-GP und GP zur mehrkriteriellen Optimierung finden zunehmend Anwendung. Zu den wichtigsten (und erfolgreichsten) Anwendungsgebieten für GP zählen Modellierung, symbolische Regression, Bild- und Signalverarbeitung, Finanzzeitreihenanalyse und ökonometrische Modelle, industrielle Prozesssteuerung, Medizin, Biologie und Bioinformatik. Beispielsweise entwickelte Kantschik [79] mit Hilfe von GP Schachprogramme. Die dabei gewonnenen Erfahrungen nutzte er zur Entwicklung eines komplexen GP-Systems, das von der *DIP GmbH* kommerziell eingesetzt wird.

Seit einiger Zeit wird die sog. *Multi-Objective Symbolic Regression* (MOSR oder auch Pareto-GP) erfolgreich als Modellierungstechnik angewendet. Basierend auf GP sollen dabei zwei widersprüchliche Zielsetzungen, einerseits Modellqualität (Prognosegüte) und andererseits Modellkomplexität, optimiert werden [96]. Pareto-GP kann zur automatischen Variablenauswahl, Berechnung sehr großer Datenmengen und Entwicklung zuverlässiger Modelle eingesetzt werden. Verfahren zur Vermeidung von Overfitting und zur Identifikation von Ausreißern werden momentan entwickelt. Der Einsatz von Pareto-GP bettet sich sehr gut in die Vorgehensweise der experimentellen Versuchsplanung (*Design of Experiments*, DoE) ein. Der Antragsteller steht im intensiven Austausch mit Forschern der

Universitäten Tilburg, Eindhoven und Antwerpen sowie mit der Arbeitsgruppe um Guido Smits (*Dow Benelux B.V.*), die diese Verfahren erfolgreich in der industriellen Praxis einsetzen.

Finanzwirtschaft In der Finanzwirtschaft werden die verschiedensten Verfahren zur Prognose von Zeitreihen eingesetzt. Neben klassischen Verfahren, wie gleitende Mittel (MA), autoregressive Schemata (AR), ARMA-Prozesse und (G)ARCH-Regressionsverfahren, werden auch CI-Methoden wie NN, EA, Fuzzy-Systeme und GP eingesetzt. Es zeigt sich hier kein einheitliches Bild, jedoch werden heute noch mehrheitlich klassische Verfahren genutzt. Diese werden meistens in Banken oder Fondgesellschaften intern implementiert. Um der Konkurrenz keinen Einblick in die eigene Arbeit zu gewähren, werden Forschungsergebnisse nicht (oder nur unvollständig und mit einer zeitlichen Verzögerung) veröffentlicht oder patentiert. Neuronale Netze werden seit den 90er Jahren immer seltener in der Finanzwirtschaft eingesetzt. Ein Grund dafür mag sein, dass NN tendenziell das Modell zu gut an die Daten anpassen und nicht flexibel auf neue Daten reagieren können (sog. „Overfitting“). Diese Systeme waren oft in-house Lösungen, so dass deren Entwickler keine Experten im Umgang mit NN waren und das Problem dadurch noch verstärkt wurde.

Im Bereich GP ist außer dem vTrader-System der *DIP GmbH*, gemäß durchgeführter Recherchen, noch kein weiteres System im Praxiseinsatz. Jedoch gibt es seit einigen Jahren vermehrt wissenschaftliche Veröffentlichungen in diesem Bereich [77, 101]. Diese basieren jedoch auf stark vereinfachten Modellannahmen, so dass diese Verfahren (noch) nicht für den Praxiseinsatz geeignet sind. Das Graphbasierte GP-System der *DIP GmbH* ist insofern einmalig, als dass es seit einigen Jahren erfolgreich in der Praxis eingesetzt wird und ohne vereinfachende Modellannahmen auskommt.

Eine interessante Übersicht bieten Brabazon u. O'Neill [66]. Sie stellen bioanaloge Algorithmen zur Modellierung ökonomischer Problemstellungen dar. Es werden Algorithmen aus den Bereichen NN, *evolutionary computation* (EC), *social systems* und *artificial immune systems* vorgestellt. Die im Rahmen des FIWA-Projektes eingesetzten GP-Verfahren werden darin erwähnt. Die Hauptfragestellung in diesem Buch befasst sich mit der Vorhersage von Finanzdaten, wobei statt eines linearen Regressionsmodells bioanaloge Verfahren zum Einsatz kommen. Neben der Modellauswahl werden die folgenden Probleme behandelt: Auswahl der Daten, Untersuchung der Ursache-Wirkung, Datenvorverarbeitung und Auswahl des Fehlermaßes.

Baranski [59] analysiert Evolutionsstrategien zur Portfoliooptimierung unter Nebenbedingungen. Diese Arbeit wurde vom Antragsteller initiiert und von Dr. Kantschik (*DIP GmbH*) betreut. Weitere Beispiele für GP-Anwendungen in der Finanzwirtschaft sind in Poli u. a. [89, S. 123ff] zu finden.

Wasserwirtschaft Im Projekt FIWA wurden auch Prognosemodelle für Abwassersysteme untersucht. Die Untersuchungen konzentrierten sich dabei auf Regenüberlaufbecken. Der aktuelle Stand der Forschung vor Projektbeginn ist im Folgenden kurz dargestellt.

Die Aufgabe von Kläranlagen besteht darin, Abwasser so zu behandeln, dass für Mensch und Natur kein Schaden entsteht. Anhand von Messungen können Simulationen und Berechnungen zum niederschlagsabhängigen Abfluss und der Beschaffenheit des Abwassers durchgeführt werden, die dann in die Planungen einbezogen werden. Zudem können Messungen zur Überwachung und Steuerung von Regenüberlaufbecken durchgeführt werden. Durch die Erfassung von Messdaten

zu Abfluss und Stofftransport, also quantitativer und qualitativer Daten in Kanalnetzen, können neue Kanäle simulativ geplant oder Szenarien für die Sanierung alter Netze durchgespielt werden. Moderne Simulatoren wie SIMBA [78] oder SWMM [98] können den zeitlichen Verlauf der Ablaufkonzentrationen und auch den zeitlichen Verlauf des Zustandes und der hydraulischen Belastung in allen Teilen der Kläranlage berechnen.

Bartz-Beielstein u. a. [63] untersuchen Vorhersagemodelle für Kläranlagen und Kanalnetze mit CI-Verfahren. Vergleichbare Arbeiten, die allerdings meist nicht den Fokus auf Steuer- und Regelungstechnik legen, sind in Fuchs u. Seggelke [72], Grüning u. Orth [74] sowie Schilling [92] veröffentlicht. Des Weiteren beschreibt Burkhart [68] die Erprobung und Optimierung von Vorklärbeckenmodellen.

Internationale Beachtung finden die von Prof. Bongards durchgeführten Forschungsprojekte zum Thema Abwassertechnik. Prof. Bongards arbeitet zusammen mit den Professoren Klasen, Scheuring, Konen und dem Antragsteller im Forschungsschwerpunkt COSA an der FH Köln. Im Forschungsvorhaben FIWA fand eine enge Zusammenarbeit mit Prof. Bongards statt, die bisher durchgeführten gemeinsamen Forschungen wurden fortgeführt und ausgeweitet. Das von Prof. Bongards geleitete *Gummersbach Environmental Computer Center* (GECO>C) führt zahlreiche Projekte im Bereich Abwassertechnik durch (siehe auch Abschnitt 2.5). Weiterführende Erläuterungen zum wasserwirtschaftlichen Stand der Technik und zur Zielsetzung einer Prognose und Regelung von Füllständen in Regenüberlaufbecken werden in Hilmer u. Bongards [76] und Grüning [73] gegeben.

Generalisierbarkeit Von großer Bedeutung ist eine effiziente Planung und Durchführung der Vergleiche von CI-basierten und klassischen Verfahren sowie die statistische Absicherung der gefundenen Ergebnisse. Die hierfür benötigten Methoden können unter dem Begriff DoE (Design of Experiments) zusammengefasst werden. Der Einsatz von DoE und modernen Erweiterungen wie *Design and Analysis of Computer Experiments* ist ein wichtiges Element, um verschiedene Modellierungen effizient zu vergleichen und ihr Potenzial zu beurteilen [91],[62]. Durch die steigende Rechnerleistung können Simulationen und auch die zugehörigen Modellverifikationen und Validierungen heutzutage bereits mit relativ geringem Hardwareaufwand durchgeführt werden, so dass die Ergebnisse einzelner Experimente (Simulationen, Prognosen) relativ schnell berechnet und überprüft werden können.

Generalisierbarkeit bezieht sich nicht nur auf eine einzelne Problemstellung, sondern versucht Gemeinsamkeiten bei mehreren Problemen zu betrachten und so zur Theoriebildung beizutragen. Eine systematische Vorgehensweise bei der Generalisierung bietet der *neue Experimentalismus*. Der neue Experimentalismus stellt eine einflussreiche wissenschaftstheoretische Richtung dar, die versucht, die Kluft zwischen Theorie und Praxis zu überwinden. Dabei kommt sorgfältig durchgeführten Experimenten eine zentrale Rolle zu. Der neue Experimentalismus erweitert dabei den von Popper entwickelten Wissenschaftsbegriff, der im Wesentlichen auf dem Gedanken der Falsifikation beruht, um Methoden der Verifikation [90]. Zu den prominenten Vertretern des neuen Experimentalismus zählen Ian Hacking, Deborah Mayo und Sir David Roxbee Cox [75, 85, 70, 71]. 2006 fand auf Initiative von Deborah Mayo das „First Symposium on Philosophy, History, and Methodology of E.R.R.O.R (Experimental Reasoning, Reliability, Objectivity and Rationality): Induction, Statistics, Modelling“ statt. Die Generalisierbarkeit experimentell gefundener Daten war ein zentrales Thema dieses Symposiums. So wird in Parker [88] Simulation aus der Perspektive des neuen Experimenta-

lismus beschrieben. Bartz-Beielstein [61] zeigt, wie Methoden des neuen Experimentalismus in der Informatik eingesetzt werden können.

Auch von Theoretikern wird der Versuch unternommen, die Kluft zur (experimentellen) Praxis zu überwinden. *Algorithm Engineering* kann hier als eine wichtige Forschungsrichtung genannt werden [69].

2.5 Zusammenarbeit mit anderen Stellen

Die Projektpartner des Projekts FIWA befinden sich auch weiterhin in engem Kontakt. Zur Abstimmung der Anwendung der entwickelten Systeme in der Finanzwirtschaft finden Gespräche mit den Projektpartnern *Quaesta Capital GmbH* und *DIP GmbH* statt. Im Rahmen dieser Zusammenarbeit wurde die GECCO INDUSTRIAL CHALLENGE 2011 zum Vergleich von Prognosesystemen für Finanzzeitreihen erfolgreich durchgeführt.⁴ Die von der Association for Computing Machinery (ACM) ausgerichtete GECCO ist die wichtigste internationale Konferenz in ihrem Fachgebiet. Prof. Dr. Bartz-Beielstein zählt zu den Organisatoren dieser Konferenz. Das Projekt FIWA war über seine gesamte Laufzeit auf der GECCO sehr präsent, u.a. durch Mitgliedschaft im Programmkomitee, Organisation des renommierten Tracks „Evolutionary Computation in Practice“ und durch Durchführung von Workshops und Tutorien. Die GECCO INDUSTRIAL CHALLENGE wurde auf der GECCO 2012 zum zweiten Mal erfolgreich durchgeführt.

Regelmäßige Treffen mit der Arbeitsgruppe GECO▷C um Prof. Dr. M. Bongards (FH Köln) versorgten das Projekt mit Praxisproblemen, Fachwissen und Daten aus der Wasserwirtschaft. Die Kooperation mit dem Lehrstuhl Computergestützte Statistik an der Fakultät Statistik der *TU Dortmund* ermöglichte den intensiven Austausch über moderne Methoden der Statistik.

Mit dem ebenfalls in der Förderlinie Ingenieur Nachwuchs 2009 geförderten Projekt SOMA (Prof. Dr. W. Konen, FH Köln) bestand über die gesamte Projektlaufzeit eine intensive Zusammenarbeit. Dies zeigt sich auch in einer Vielzahl gemeinsamer wissenschaftlicher Veröffentlichungen (siehe auch Abschnitt 5). Die in der Projektlaufzeit von Prof. Dr. Bartz-Beielstein initiierte und gemeinsam mit Prof. Konen geleitete Forschungsstelle „Computational Intelligence, Optimization & Data Mining (CIOP)“ an der FH Köln gibt dieser erfolgreichen Zusammenarbeit einen institutionellen Rahmen.⁵

In der während der Projektlaufzeit initiierten und durch den Projektleiter herausgegebenen Schriftenreihe „CIOP Reports“ (ISSN, 2191-365X) erscheinen regelmäßig technische Berichte (siehe Abschnitt 5). Ausserdem besteht ein lebhafter Austausch mit regionalen Industrieunternehmen und auch international renommierten Forschern.

3 | Wissenschaftlich-technische Ergebnisse

Ziel des Projekts FIWA war die Entwicklung modularer Systeme zur Analyse und Prognose von Daten aus der Finanz- und Wasserwirtschaft mittels Verfahren der Computational Intelligence (CI) mit einem Schwerpunkt auf Genetic Programming (GP). Die eingehende Darstellung der wissenschaftlich-technischen Ergebnisse gliedert sich nach der Modul- und Arbeitspaketplanung der Vorhabensbe-

⁴siehe <http://www.sigev.org/gecco-2011/industrial-challenge.html>

⁵Eine ausführlichere Darstellung der Aktivitäten der Forschungsstelle CIOP bietet die Website <http://gociop.de>.

schreibung. Für die Arbeitspakete wurden die folgenden Kürzel verwendet: GP: Genetic Programming, F: Finanzwirtschaft, W: Wasserwirtschaft, G: Generalisierbarkeit, MT: Masterarbeiten und CS: Case-Studies.

3.1 Modul M-1 „Genetic Programming“

Die Arbeitspakete des Moduls M-1 befassten sich mit der Verallgemeinerung und Dokumentation des vTraderGP-Systems. Im Zuge dieser Verallgemeinerung wurden wesentliche Funktionen des System in die freie Statistiksoftware R portiert und unter dem Namen RGP als Open Source Software veröffentlicht. Des Weiteren wurden allgemeine Zeitreihenanalysefunktionen aus R in RGP integriert.

3.1.1 GP-1: Verallgemeinerung und Dokumentation des GP-Systems

Die *DIP GmbH* stellte dem Projekt FIWA ein professionelles GP-System (vTrader) zur Verfügung. Die für FIWA wesentlichen Funktionen dieses Systems wurden in die Statistik-Softwareumgebung R⁶ portiert und sind nun einer weiten Nutzergemeinschaft unentgeltlich als Open Source Software unter dem Namen RGP verfügbar.⁷ [39] Im Zuge der Portierung des Systems nach R wurden bestehende GP-Funktionen für den Einsatz in weiteren Anwendungsgebieten wie der Wasserwirtschaft verallgemeinert. Die Funktionalität und Modularität des Systems wurde stark erweitert und dessen Basisarchitektur verallgemeinert. Außerdem wurden alle Funktionen dokumentiert, wodurch ein effektiver Einsatz des Systems zur Bearbeitung von Prognosemodellen in der Wasserwirtschaft (W-1 und W-2) und die Nutzung durch weitere Anwender ermöglicht wird. Die Dokumentation ist als Teil des RGP Pakets als Online-Hilfe verfügbar, weitere Dokumentation steht auf der RGP Website zur Verfügung. Im November 2012 hatte RGP 29 registrierte Benutzer⁸.

RGP ist inzwischen ein leistungsfähiges Werkzeug zur Erforschung von Zeitreihenregressions- und Analyseaufgaben mit GP. Gleichzeitig ist das System modular aufgebaut, um einfach auf neue Anwendungsgebiete erweitert zu werden. Es folgt eine Übersicht der in RGP implementierten Funktionalität.

Repräsentation von GP-Individuen RGP stellt GP-Individuen als Ausdrücke der Programmiersprache R dar, so dass diese direkt vom R-Interpreter ausgewertet werden können. Dadurch kann das gesamte Spektrum der in R verfügbaren Funktionen als Bausteine für GP genutzt werden. Da R-Ausdrücke intern als Baum-Datenstrukturen dargestellt werden, kann RGP als Baum-basiertes GP-System angesehen werden. Allerdings kann die Individuen-Repräsentation bei Bedarf ersetzt werden, gemeinsam mit den assoziierten Operatoren für Initialisierung, Variation und Auswertung. In RGP steht mit *Linear GP* bereits eine alternative Darstellung zur Verfügung [94].

Neben klassischem ungetyptem GP unterstützt RGP streng typisiertes GP auf Basis des *Simply Typed Lambda Calculus* [60]. Eine besondere Eigenschaft der streng typisierten RGP Individuen-Repräsentation ist die Unterstützung von *Function Defining Subtrees*, d.h. von Lambda-Abstraktionen

⁶R ist eine freie Statistik-Softwareumgebung, die in den letzten Jahren grosse Verbreitung in der Wissenschaft und in der Finanz- und Biotechnologiebranche gefunden hat. Für Details siehe <http://www.r-project.org/>.

⁷RGP steht unter <http://rsymbolic.org/> zum Download zur Verfügung.

⁸Quelle: Anzahl der aktiven registrierten Benutzer auf der RGP Website <http://rsymbolic.org/>

(anonymen Funktionsausdrücken). In Verbindung mit dem RGP Typ-System, welches Funktionstypen unterstützt, ermöglicht dies die Nutzung von Funktionen höherer Ordnung, z.B. *folds*, *mappings* und *convolutions*, als Bausteine für GP. Diese Funktionen sind insbesondere für die Zeitreihenprognose nützlich.

RGP beinhaltet einen regelbasierten Übersetzer für die Transformation von R-Ausdrücken. Dieser Mechanismus kann zur Vereinfachung von GP-Individuen genutzt werden um diese dem Benutzer verständlicher präsentieren zu können. Ein weiteres Anwendungsgebiet ist die Vermeidung von *Bloat*, d.h. dem unkontrollierten Zuwachs der Individuen-Komplexität in einem GP-Lauf.

GP-Operatoren RGP implementiert diverse Standardoperatoren für Initialisierung, Variation und Selektion. [80] Die Initialisierung von GP-Individuen kann durch die konventionellen *grow*- und *full*-Strategien der Baumstruktur-Erstellung geschehen. Bei der Nutzung von streng typisiertem GP beachten die implementierten Algorithmen Typ-Nebenbedingungen und erstellen nur Typ-korrekte R-Ausdrücke. Die Strategien der Initialisierung können frei kombiniert werden, um z.B. die verbreitete *ramped-half-and-half*-Strategie umzusetzen. RGP implementiert klassisches wie auch streng typisiertes Crossover. Zusätzlich stehen mehrere klassische und streng-typisierte Mutationsoperatoren zur Verfügung. Sowohl Reihenfolge und Art der Variation ist frei konfigurierbar. Das System bietet mehrere Strategien für die Individuenselektion. Für die Unterstützung mehrkriterieller Selektion stehen Komplexitätsmasse für GP-Individuen zur Verfügung. Mehrere evolutionäre Algorithmen werden als Suchstrategien unterstützt, z.B. zeigt Abbildung 2 verteiltes Cluster-basiertes Niching [31].

Testfunktionen Zur visuellen Validierung der korrekten Funktion einer RGP-Installation sind mehrere einfache Testfunktionen für die symbolische Regression durch GP im RGP-Paket enthalten. Abbildung 3 zeigt eine Auswahl dieser Funktionen. Die wahren Funktionswerte sind durchgängig schwarz, die durch RGP erzeugten Lösungen schwarz gestrichelt dargestellt.

3.1.2 GP-2: Entwicklung weiterer allgemeiner Zeitreihenanalysefunktionen

Einer der Hauptfaktoren für die in GP-1 beschriebene Portierung des vTraderGP-Systems nach R war die große Anzahl und hohe Qualität der bereits in R frei verfügbaren allgemeinen Zeitreihenanalysefunktionen. Diese Funktionen sind nun sehr einfach direkt durch streng typisiertes GP nutzbar.

Technisch stellt RGP Individuen für Zeitreihenprognosemodelle in der streng typisierten funktionalen Untermenge `gpL` der Programmiersprache R dar. Die Menge der Datentypen T ist wie folgt definiert: Ein Typ $\tau \in T$ ist entweder ein R-Basistyp aus $B := \{\text{logical, integer, numeric, string, date_time, time_series}\}$, oder ein n -ärer Funktionstyp $(\delta_1, \dots, \delta_n) \rightarrow \nu$, für Basistypen $\delta_1, \dots, \delta_n$, und ν . Damit ist die Menge der Typen T vollständig definiert.

Die Menge der `gpL`-Ausdrücke (Terme) ist nun wie folgt definiert: Ein `gpL`-Term ist entweder ein Literal c des Basistyps $\beta \in B$, oder eine Eingabevariable x des Basistypes $\beta \in B$, oder eine n -äre Funktionsanwendung $f(t_1, \dots, t_n)$ des Typs $\tau \in T$, wobei f vom Typ $(\delta_1, \dots, \delta_n) \rightarrow \tau$ und die Terme t_i von Typen $\delta_i \in B$ für $i \in [1, n]$ sein müssen. Damit ist die Menge der `gpL`-Terme vollständig definiert.

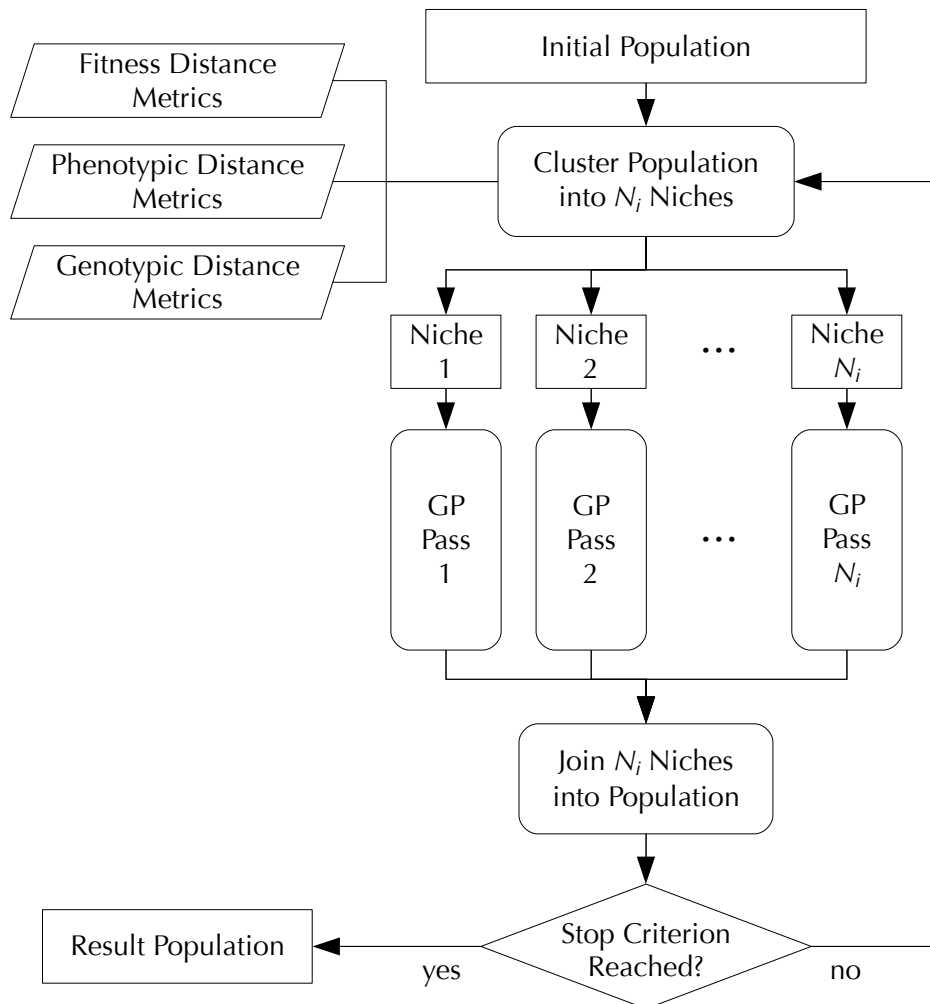


Abbildung 2: Ablaufdiagramm des verteilten Cluster-basierten Niching Algorithmus in RGP: Die initiale Population von GP-Individuen wird in N_i Nischen aufgeteilt, basierend auf Fitness-Distanz, phänotypischer Distanz, oder auf genotypischer Distanz. Die Nischen werden auf einen Rechencluster verteilt, wobei jeder Rechenknoten einen unabhängigen GP-Pass (GP-Lauf) durchführt. Die Ergebnisse dieser Läufe werden zu einer einzelnen Ergebnis-Population zusammengeführt. Dieser Prozess wiederholt sich, bis ein Stop-Kriterium erfüllt ist.

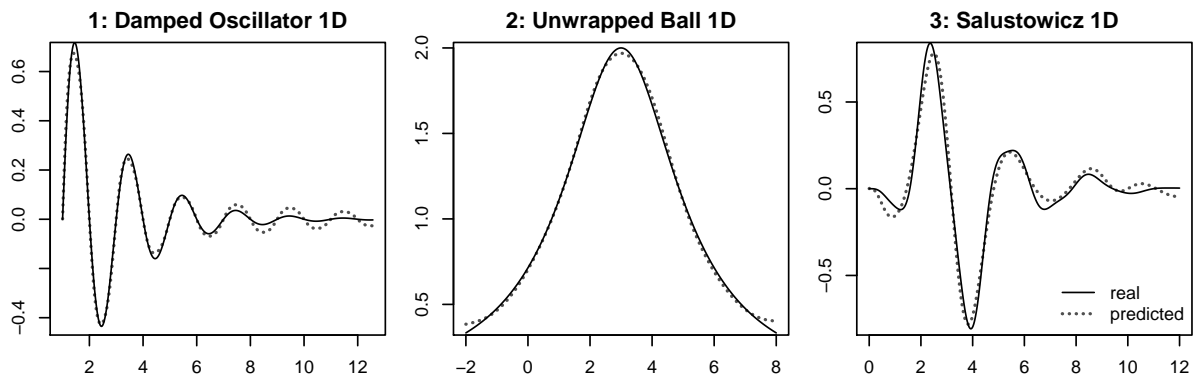


Abbildung 3: Plots univariater Testfunktionen für symbolische Regression durch GP, wie in RGP enthalten: Die wahren Funktionswerte sind durchgängig schwarz, die durch RGP erzeugten Lösungen schwarz gestrichelt dargestellt.

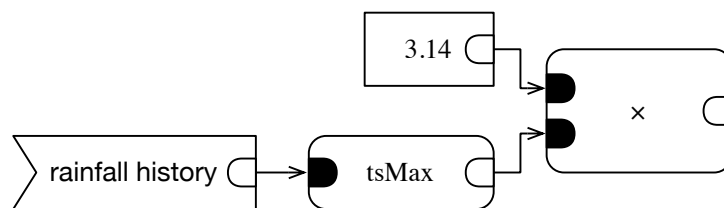


Abbildung 4: Ein einfaches Beispiel eines GP-Individuums, dargestellt als Baum-Datenstruktur, wie sie in RGP verwendet wird.

Als Beispiel soll der gültige gpL-Term $3.14 \times \text{tsMax}(\text{rainfall history})$ dienen. Dieser Term ist eine Funktionsanwendung und hat den Basistyp `numeric`. Die Funktion \times hat den Funktionstyp `(numeric,numeric) → numeric`, die Funktion `tsMax` den Funktionstyp `(time_series) → numeric`, das Literal `3.14` den Basistyp `numeric` und die Eingabevariable `rainfall history` den Basistyp `time_series`. Abbildung 4 zeigt, wie dieser Term in der RGP Baum-Datenstruktur dargestellt wird.

Um RGP für die Zeitreihenprognose einzusetzen, werden GP-Individuen des Typs `(time_series,..., time_series) → numeric` evolviert. Diese GP-Individuen werden als Zeitreihenprognosemodelle verwendet, indem sie als Funktionen auf ein Fenster der Eingabe-Zeitreihen angewandt werden. Der resultierende Wert des Basistyps `numeric` wird als Prognose interpretiert. Der Prognosehorizont, wie auch die zu prognostizierende Zeitreihe, ist in der Fitnessfunktion wie folgt kodiert: Als Trainingsmenge wird eine Zeitreihe historischer Daten von S_{target} , der zu prognostizierenden Zeitreihe, wie auch von $\mathcal{S} := \{S_{input_1}, \dots, S_{input_n}\}$, der Menge der Eingabe-Zeitreihen verwandt (S_{target} kann in \mathcal{S} enthalten sein). Die Prognosegenauigkeit eines Individuums I zum Zeitpunkt t mit Prognosehorizont h wird wie folgt bestimmt: Alle Eingabe-Zeitreihen werden auf das Fenster $[t - (h+w), t - h]$ beschnitten, hierbei ist w eine vorab gewählte Fensterlänge. Dann wird I auf die Menge der beschnittenen Eingabe-Zeitreihen angewandt, um die Prognose von I für $S_{target}(t)$ zu berechnen. Dann wird der quadratische Fehler (Squared Error, SE) zwischen dem wahren Wert der Zeitreihe $S_{target}(t)$ und dieser Prognose bestimmt. Der Root Mean Squared Error der Prognose wird approximiert, indem die Wurzel von einer großen Zahl von zufällig gezogenen SE-Werten der Trainingsmenge gezogen wird.

Diese Approximation wird als Fitness des GP-Individuums I genutzt.

In der Praxis kann ein GP-Individuum I als Prognosemodell eingesetzt werden, indem es als R-Funktion auf ein Fenster der Länge w der Eingabe-Zeitreihen angewendet wird.

3.2 Modul M-2 „Finanzwirtschaft“

Ziel der Arbeitspakete im Modul M-2 war die Entwicklung von Zeitreihenanalyse- und Prognosefunktionen für Anwendungen in der Finanzwirtschaft und deren Integration in das GP-System. Ein weiteres Ziel war die Optimierung der Prognosemodellauswertung, um die Effizienz des GP-Systems bei der Evolution von Handelssystemen zu steigern.

3.2.1 F-1: Entwicklung von auf Finanzzeitreihen spezialisierten Analysefunktionen

Da sich die Statistik-Softwareumgebung R in den letzten Jahren in der Finanzwirtschaft immer stärker verbreitet hat, stehen dem Projekt durch die Portierung des vTraderGP-Systems in R bereits eine große Anzahl auf Finanzzeitreihen spezialisierter Analysefunktionen in hoher Qualität zur Verfügung. Diese Funktionen sind in RGP als GP-Bausteine nutzbar.⁹

3.2.2 F-2: Optimierung der Prognosemodellauswertung

Der Trading-Simulator der *DIP GmbH*, welcher den Handel mit Finanzinstrumenten unter realistischen Bedingungen simuliert, um die Prognosegenauigkeit anhand historischer Finanzzeitreihen realistisch bewerten zu können, wurde nach R portiert. Die Effizienz des Simulators wurde im Rahmen dieser Portierung optimiert, wodurch die Gesamteffizienz der Anwendung von GP für die Finanzzeitreihenprognose gesteigert werden konnte. Gleichzeitig wurde der Trading Simulator nun als Open Source Software zur Verfügung gestellt und ist eigenständig nutzbar, was den Vergleich verschiedener Prognoseverfahren, auch über GP hinaus, sehr vereinfacht.

Der Trading-Simulator kam bereits sehr erfolgreich zur Evaluation der Prognosegenauigkeit verschiedener CI-basierter Verfahren zur Finanzzeitreihenprognose zum Einsatz. Abbildung 5 zeigt ein Ablaufdiagramm zum Einsatz des Simulators.

Abbildung 6 zeigt die akkumulierte Rendite beim Devisenhandel mit dem Währungspaar AUDUSD. Daten dieser Form werden durch den Trading-Simulator verarbeitet.

3.2.3 F-3: Entwicklung eines Bewertungswerkzeugs für Finanzzeitreihen-Prognosemodelle

Im Rahmen dieses Arbeitspaketes wurden graphische und statistische Werkzeuge zur Bewertung der vom GP-System gelieferten Prognosemodelle anhand verschiedener Kriterien erstellt. Dazu wurden vorhandene und frei verfügbare graphische und statistische Werkzeuge der Softwareumgebung R auf das Problem angepasst, wodurch der Aufwand gegenüber einer Neuentwicklung verringert wurde. Zusätzlich können Anwender aus der Finanzwirtschaft die für sie gewohnten R-basierten Werkzeuge zur Modellbewertung weiter nutzen.

⁹Siehe beispielsweise <http://cran.r-project.org/web/views/Econometrics.html> für eine aktuelle Übersicht zu Finanzzeitreihen-Analysefunktionen für R.

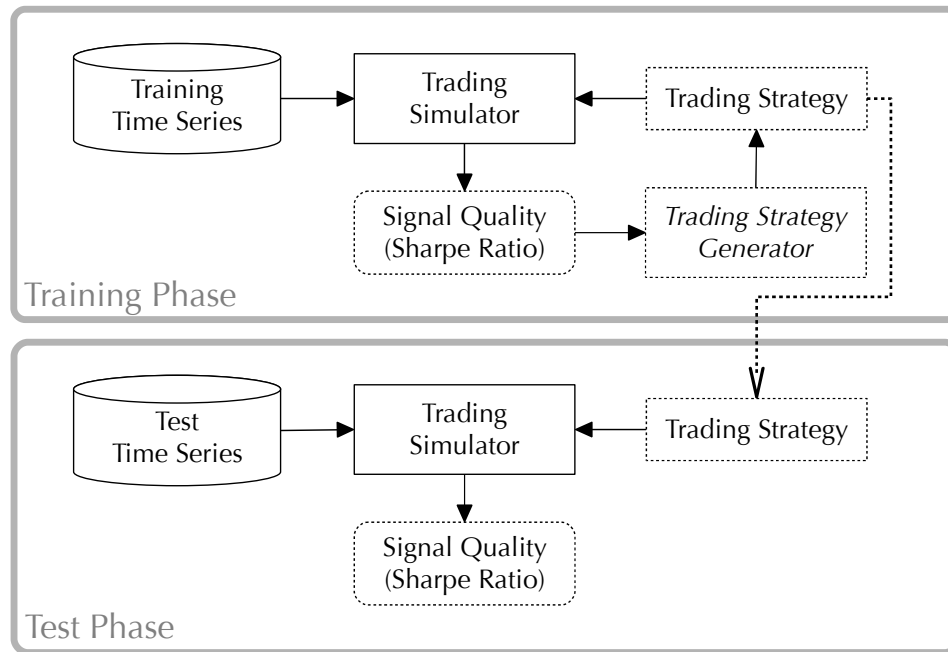


Abbildung 5: Ablaufdiagramm zum Einsatz des im Projekt FIWA entwickelten Trading-Simulators: Während eines GP-Laufs, also in der Training-Phase, wird die Qualität der durch GP als *Trading Strategy Generator* erzeugten Handelsstrategien durch simulierten Handel auf Basis von historischen Daten der Trainings-Zeitreihen geschätzt. In einer nachfolgenden unabhängigen Test-Phase wird die Qualität der durch GP erzeugten Handelsstrategien durch simulierten Handel auf Test-Zeitreihen validiert.

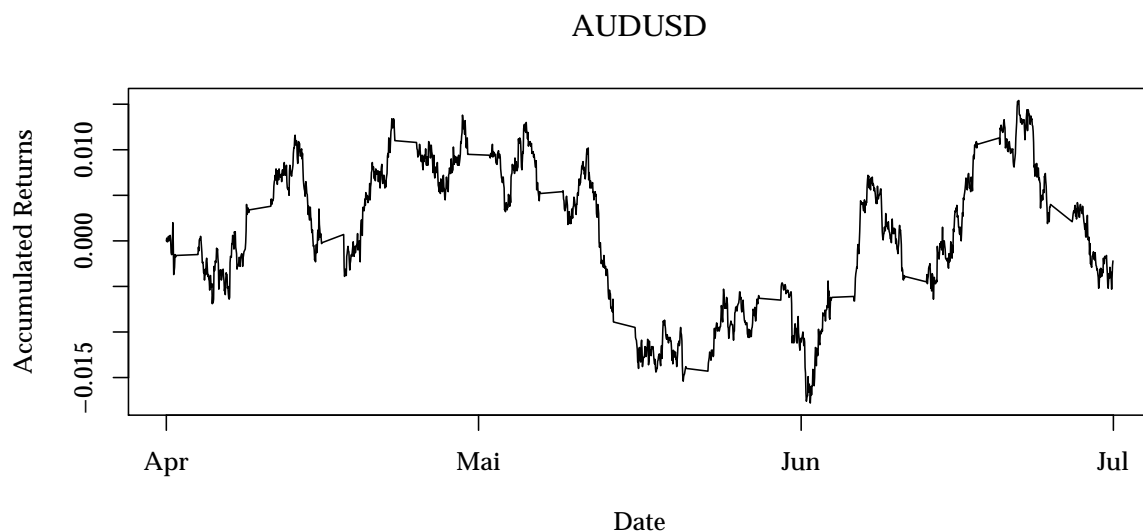


Abbildung 6: Akkumulierte Rendite beim Devisenhandel mit dem Währungspaar AUDUSD. Daten dieser Form werden durch den im Projekt FIWA entwickelten Trading-Simulator zur Schätzung der Genauigkeit von Prognosemodellen für Finanzzeitreihen genutzt.

In Verbindung mit der im Arbeitspaket F-2 erstellten optimierten Prognosemodellauswertung bildeten die in diesem Arbeitspaket entwickelten Werkzeuge eine ideale Basis für die Ausrichtung eines internationalen Wettbewerbs zum Vergleich von Prognosesystemen für die Finanzwirtschaft unter realistischen Bedingungen. Abschnitt 3.6 und die Website <http://gociop.de/gecco-industrial-challenge/gecco-2011-industrial-challenge/> enthalten Details zu diesem Wettbewerb, der GECCO INDUSTRIAL CHALLENGE 2011.

Abbildung 7 zeigt das im Rahmen des Projekts weiterentwickelte Werkzeug zur graphischen Darstellung, Manipulation und Auswertung von Prognosemodellen in der Finanzwirtschaft. Gezeigt wird die graphische Analyse des Euro/Cable-Indikators, eines Indikators der technischen Finanzmarktanalyse. Diese Indikatoren liefern unter anderem Eingabedaten für die im Projekt entwickelten Prognosemodelle.

3.2.4 F-4: Evaluation der Prognosegenauigkeit für Finanzzeitreihen

Die Evaluation der erreichten Prognosegenauigkeit für Finanzzeitreihen mit Hilfe der im Arbeitspaket F-3 entwickelten Werkzeuge konnte im Rahmen der GECCO INDUSTRIAL CHALLENGE 2011 in erweitertem Umfang erfolgen. Die Ergebnisse dieser Analysen werden ein Kernkapitel der im Projektrahmen erstellten Dissertation bilden. Dabei erfolgte nicht nur ein Vergleich mit klassischen Methoden der Zeitreihenprognose, sondern auch mit hochaktuellen Methoden aus der CI.

Abbildung 8 zeigt eine Übersicht der Ergebnisse der ersten Runde der GECCO INDUSTRIAL CHALLENGE 2011. Ziel des Wettbewerbs war die Entwicklung von erfolgreichen CI-basierten algorithmischen Strategien für den Devisenhandel, beispielsweise mit GP. Die Berechnung des Qualitätsmasses (Sharpe Ratio) erfolgte mit Hilfe des im Arbeitspaket F-2 erstellen Trading-Simulators. Zum Vergleich ist die gemessene Qualität einer einfachen *Buy and Hold*-Strategie, sowie die Qualität einer einfachen vom Projektpartner *Quaesta Capital GmbH* gelieferten Basis-Strategie (*Simple Proprietary*) ebenfalls dargestellt.

Abbildung 9 zeigt das Verhalten einiger der zur GECCO INDUSTRIAL CHALLENGE 2011 eingereichten Handelsstrategien im Vergleich zur einer einfachen *Buy and Hold*-Strategie. Als Datenbasis dienten Renditen des Währungspaares EURUSD. Die gezeigten Strategien erzielten relativ zur *Buy and Hold*-Strategie einen höheren Gewinn bei geringerem Risiko.

3.2.5 CS-1: Vergleichende Übersicht: Prognosemodelle in der Finanzwirtschaft

Diese Case Study wurde als Masterarbeit MT-1 mit dem Titel „Algorithmic Stock Trading as a Real World Test Problem for Time Series Prediction Methods“ vergeben und Anfang 2011 erfolgreich fertiggestellt.

3.2.6 CS-2: Vergleich des FIWA-GP Prognosemodells mit anderen Modellen in der Finanzwirtschaft

Diese Case Study wurde in erweitertem Umfang als GECCO INDUSTRIAL CHALLENGE 2011 durchgeführt und erfolgreich abgeschlossen. Wie bereits unter F-4 erwähnt, wird die Ergebnisanalyse ein Kernkapitel der im Projekt erstellten Dissertation bilden. Zu dieser Analyse werden Methoden der experimentellen Versuchsplanung und des neuen Experimentalismus (wie z.B. SPO) herangezogen.

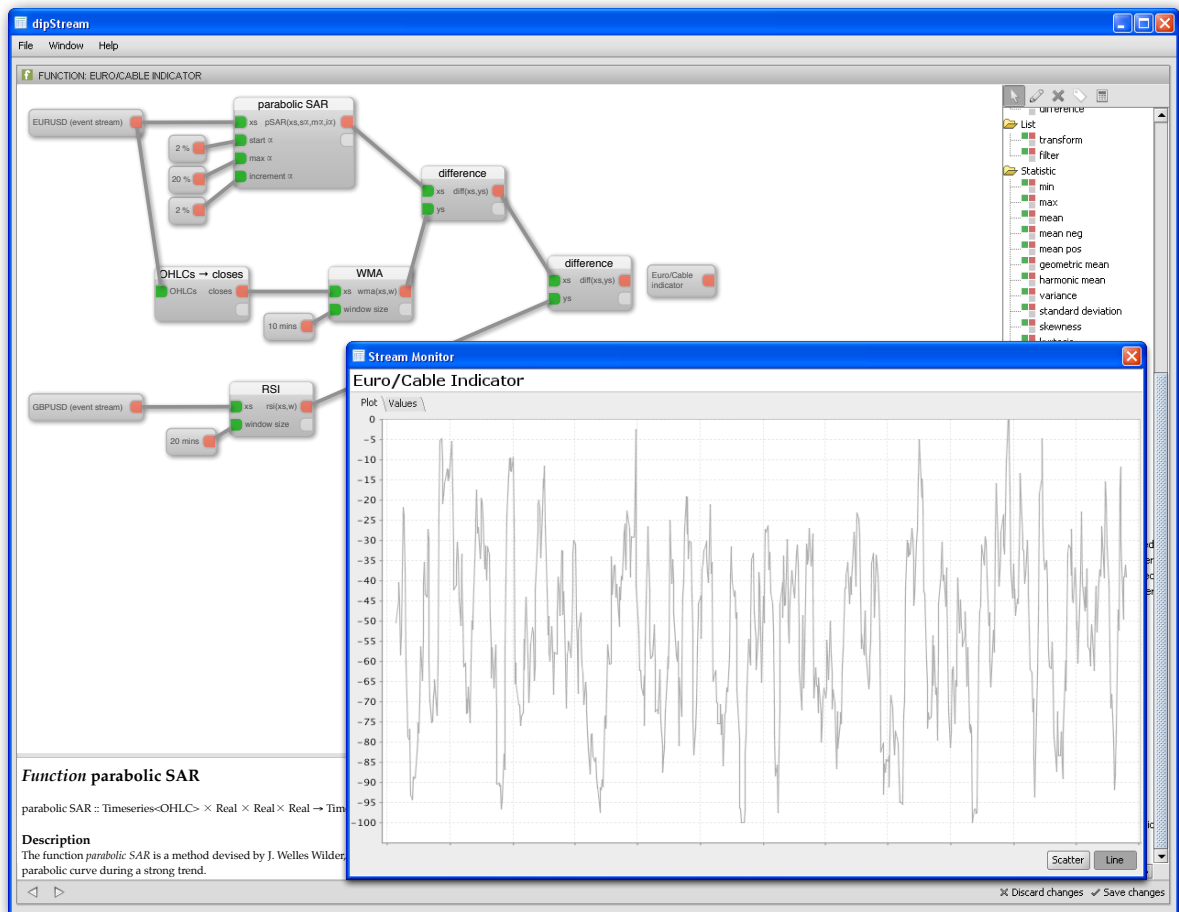


Abbildung 7: Werkzeug zur graphischen Darstellung, Manipulation und Auswertung von Prognosemodellen in der Finanzwirtschaft. Gezeigt wird die graphische Analyse des Euro/Cable-Indikators, eines Indikators der technischen Finanzmarktanalyse. Diese Indikatoren liefern unter anderem Eingabedaten für die im Projekt FIWA entwickelten Prognosemodelle.

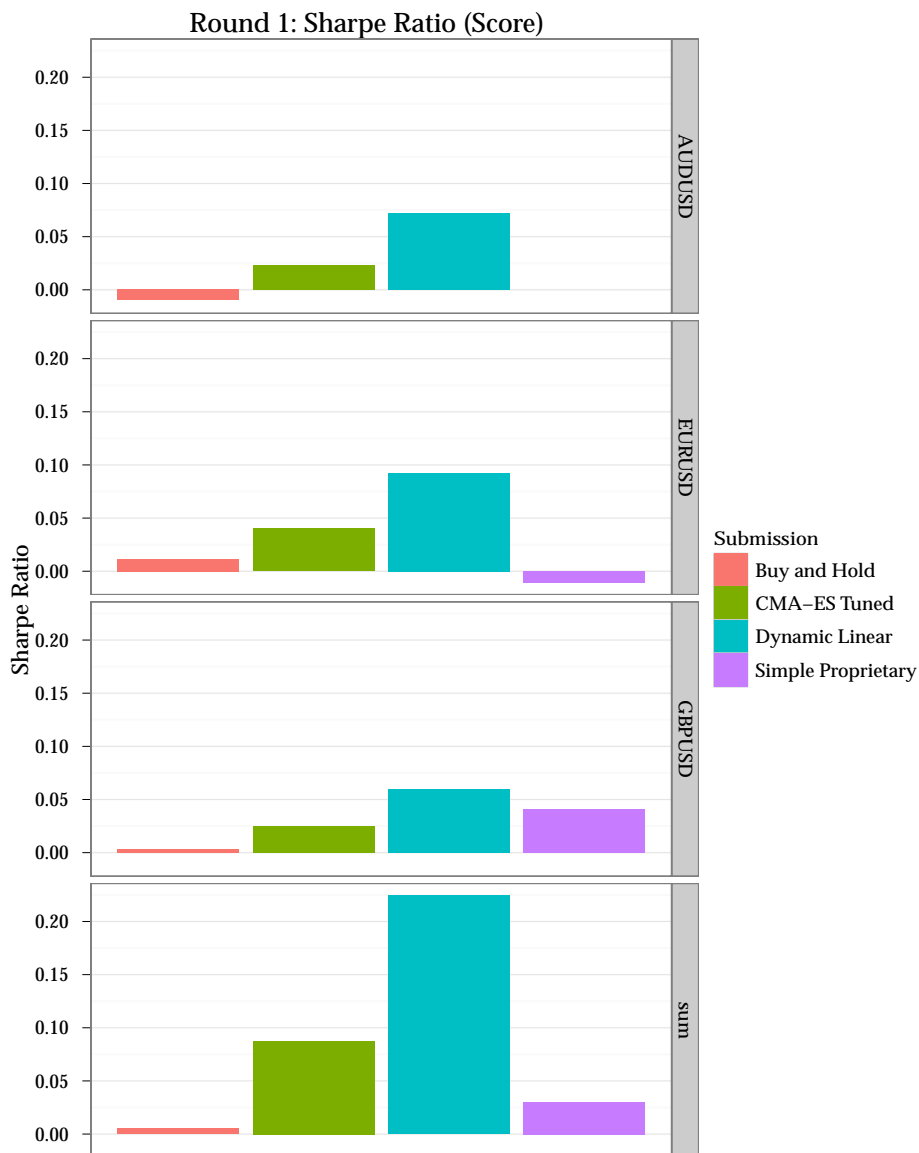


Abbildung 8: Zusammenfassung der Ergebnisse der ersten Runde der GECCO INDUSTRIAL CHALLENGE 2011. Ziel war die Entwicklung von erfolgreichen CI-basierten algorithmischen Strategien für den Devisenhandel, beispielsweise mit GP. Die Berechnung des Qualitätsmasses (Sharpe Ratio) erfolgte mit Hilfe des im Arbeitspaket F-2 erstellten Trading-Simulators. Zum Vergleich ist die gemessene Qualität einer einfachen *Buy and Hold*-Strategie, sowie die Qualität einer einfachen vom Projektpartner *Quaesta Capital GmbH* gelieferten Basis-Strategie (*Simple Proprietary*) ebenfalls dargestellt.

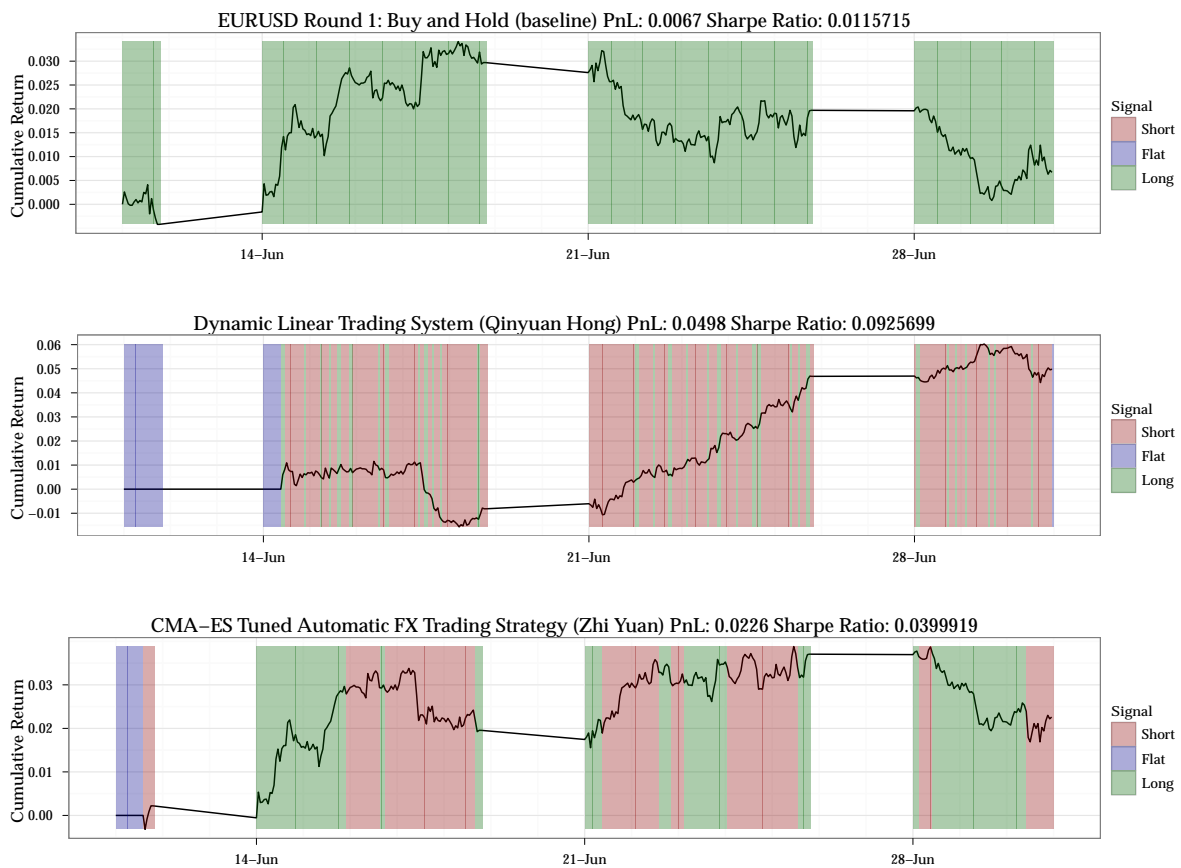


Abbildung 9: Verhalten einiger der zur GECCO INDUSTRIAL CHALLENGE 2011 eingereichten CI-basierten Handelsstrategien im Vergleich zur einer einfachen *Buy and Hold*-Strategie. Datenbasis waren Renditen des Währungspaares EURUSD. Die gezeigten Strategien erzielten relativ zur *Buy and Hold*-Strategie einen höheren Gewinn bei geringerem Risiko.

3.2.7 MT-1: Erstellung von Softwarepaketen für Finanzzeitreihen

Wie bereits im Zwischenbericht 2010 erwähnt, konnte der Aufwand für Eigenentwicklungen von Softwarepaketen für die Analyse von Finanzzeitreihen durch die Nutzung der Softwareumgebung R soweit gesenkt werden, dass eine Vergabe dieser Arbeiten als Masterarbeit nicht mehr nötig war. Wie bereits unter CS-2 beschrieben, ermöglichte dies die Vergabe der in der Vorhabensbeschreibung angekündigte Case Study CS-1 als Masterarbeit MT-1. Die Arbeit mit dem Titel „Algorithmic Stock Trading as a Real World Test Problem for Time Series Prediction Methods“ vergleicht vorhandene Modelle für die Finanzzeitreihenprognose aus der klassischen Statistik und aus dem Bereich CI und wurde Anfang 2011 erfolgreich fertiggestellt.

3.3 Modul M-3 „Wasserwirtschaft“

Im Modul M-3 wurden Daten und Testanwendungen aus der Wasserwirtschaft erfasst und dokumentiert. Außerdem wurden auf die Wasserwirtschaft angepasste Analysefunktionen entwickelt und in das GP-System RGP integriert. Ein weiteres Ergebnis stellt die erfolgreiche Entwicklung graphischer Werkzeuge für die Bewertung und Integration von Wasserwirtschaft-Prognosemodellen dar.

3.3.1 W-1: Datenerfassung und GP-Anpassung für Daten aus der Wasserwirtschaft

Das Arbeitspaket W-1 befasste sich mit der Erfassung von Daten aus der Wasserwirtschaft. Dazu gehört die Aufnahme der Daten selbst, ihre Dokumentation und ihre Aufbereitung. Die Aufbereitung soll die Daten in einem standardisierten Format für die Analyse mit im Projekt eingesetzten Methoden zu Verfügung stellen. Zur Datenaufnahme gehörte dabei auch die Zusammenstellung von Problemstellungen die mit den Daten verbunden sind.

Ein Anwendungsbereich dafür ist die Modellprädiktive Regelung von Füllständen in Regenüberlaufbecken, wie in Abbildung 11 zu sehen. Regenüberlaufbecken fangen Belastungsspitzen von von Kanalnetzen auf, um so ein Übertritt von Abwasser in die Umwelt bestmöglich zu verhindern. Die Vorhersage von Füllständen soll eine gezieltere Steuerung des Abflusses aus den Überlaufbecken erlauben. Zu diesem Zweck wurden Daten über die Füllstände selbst und Regenmengen aufgenommen. Die Daten über Regenmengen stammen von einem Niederschlagsmesser, wie in Abbildung 10 gezeigt. Abbildung 12 zeigt beispielhaft welche Daten für diese Applikation entstehen.



Abbildung 10: Niederschlagsmessstation.



Abbildung 11: Leeres und leicht gefülltes Überlaufbecken.

3.3.2 W-2: Entwicklung von auf Wasserwirtschaftszeitreihen angepassten Analysefunktionen

Hier wurden verschiedene Analysemethoden entwickelt und verglichen, die für Wasserwirtschaftszeitreihen geeignet sind.

Ein in diesem Paket entwickeltes GP-basiertes Verfahren zur Vorhersage der Füllstände eines Regenüberlaufbeckens wird durch Flasch u. a. [37] beschrieben. Dabei konnte GP als alleinstehendes Modellierungswerkzeug keine geeigneten Vorhersagen generieren.¹⁰ Als neue Entwicklung im Projekt FIWA wurde GP eingesetzt, um die Kernel Funktionen des INT2 Modells zu optimieren. Dieses neue hybride Verfahren aus analytischen Methoden auf dem Stand der Technik (INT2) und modernen CI-basierten Methoden (GP) konnte sehr gute Vorhersagen generieren. Ein Beispiel für die Unterschiede zwischen dem klassischen INT2 Modell und dem neuen hybriden Verfahren wird in Abbildung 13 gezeigt. Für das hybride GP/INT2 System wurde mit SPO eine Optimierung der Parameter vorgenommen. Dies ermöglichte eine zusätzliche Steigerungen der Vorhersagegenauigkeit. Für den entwickelten Ansatz ergibt sich damit das in Abbildung 14 abgebildete Schema. Auf der untersten Ebene benutzt das INT2 Modell die Regen- und Füllstandsdaten um auf Basis der eingestellten Kernelfunktionen die Füllstände vorherzusagen. Die vorhergesagten Füllstände werden mit den echten verglichen, um den Fehler zu bestimmen. Dieser Fehler wird im GP-System benutzt, um die Qualität eines Satzes von Kernelfunktionen zu begutachten. Der Fehler des besten von GP gefundenen Satzes wird wiederum von SPO genutzt, um die Güte einer Konfiguration von GP-Parametern zu bewerten und zu optimieren. Der Benutzer muss dabei nur selektieren, welche Parameter (z.B.: Populationsgrößen, Mutationsraten, Selektionsdruck) optimiert werden sollen.

Am Beispiel von Support Vector Machines (SVM) für die Vorhersage von Überlaufbeckenfüllständen zeigten Koch u. a. [45], wie, neben der Datenanalyse selbst, auch die Datenvorverarbeitung mit SPO optimiert werden kann. Mit dem optimierten SVM-Ansatz konnten deutlich bessere Ergebnisse als mit dem klassischen INT2 Modell erreicht werden.

Des Weiteren führten Flasch u. a. [30] einen Vergleich von verschiedenen Kombinationen aus Parameter-Optimierung und Prognosealgorithmen durch. Hier standen neben einem mit SPO optimierten, hybriden GP-System auch weitere CI-Verfahren wie Genetische Algorithmen und Neuronale Netze zum Vergleich. Auch in diesem Fall konnte das hybride GP/INT2 Modell die anderen Metho-

¹⁰Das auf klassischen Verfahren basierende INT2 Modell ist in [47] ausführlich beschrieben.

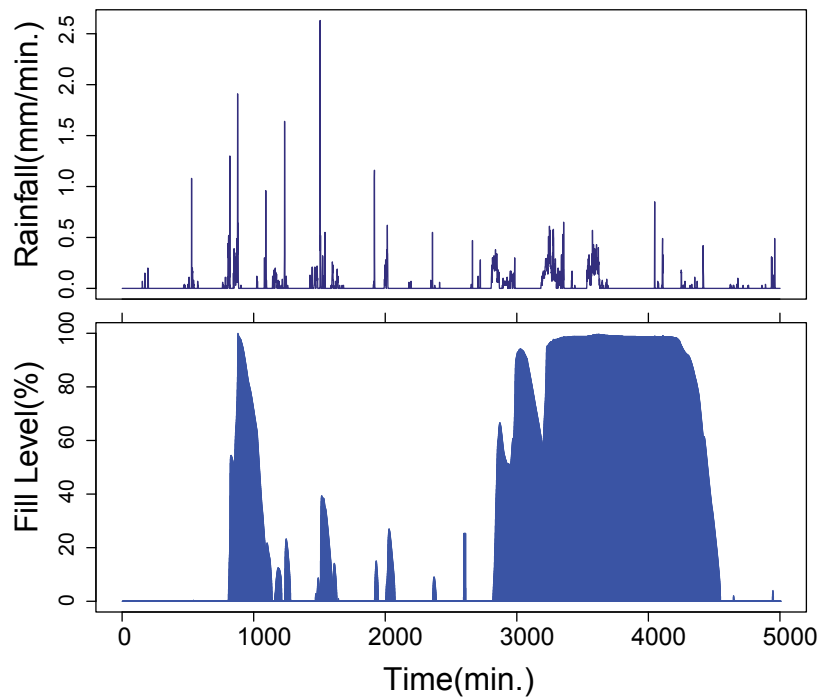


Abbildung 12: Daten über Regenmengen (oben) und Füllstände des Überlaufbeckens (unten).

den in Hinsicht auf die Vorhersagegenauigkeit klar schlagen. Die Optimierung mit SPO zeigte eine signifikante Verbesserung. Dabei war neben einer Verbesserung der mittleren Vorhersagegenauigkeit vor allem eine deutlich bessere, also kleinere, Standardabweichung zu beobachten.

3.3.3 W-3: Entwicklung eines Bewertungswerkzeugs für Wasserwirtschafts-Prognosemodelle

Analog zum Arbeitspaket F-3 wurden graphische und statistische Werkzeuge aus der Softwareumgebung R für die Bewertung der durch GP erzeugten Prognosemodelle angepasst. Zudem konnten die bereits in Arbeitspaket F-3 benutzten Tools aus der Finanzwirtschaft erfolgreich für Probleme in der Wasserwirtschaft angewendet werden. Die Ergebnisse kamen in der Folge auch im Rahmen der GECCO Industrial Challenge 2012 [40] zum Einsatz. Mit diesen Tools erstellte Visualisierungen eignen sich gut, um die Qualität von Vorhersagen einschätzen zu können, wie Abbildung 13 beispielhaft zeigt. Ein weiteres entwickeltes Tool zum interaktiven Umgang mit Zeitreihen zeigt Abbildung 15.

3.3.4 W-4: Evaluation der erreichten Prognosegenauigkeit für Wasserwirtschaftszeitreihen

Die Prognosegenauigkeit für Zeitreihen der Wasserwirtschaft wurde mit Hilfe der im Arbeitspaket W-3 entwickelten Werkzeuge evaluiert. Dabei unterstützen die Tools vor allem die visuelle Überprüfung der Prognosen. Für die Resultate der in Abschnitt 3.3.2 beschriebenen Arbeiten kamen die in W-4 erstellten Methoden und Werkzeuge für die Bewertung zur Anwendung.

Als etablierte Methoden wurden hier folgende Modellierungsansätze evaluiert und verglichen:

- Nonlinear AutoRegressive neural networks with eXogenous inputs (NARX)

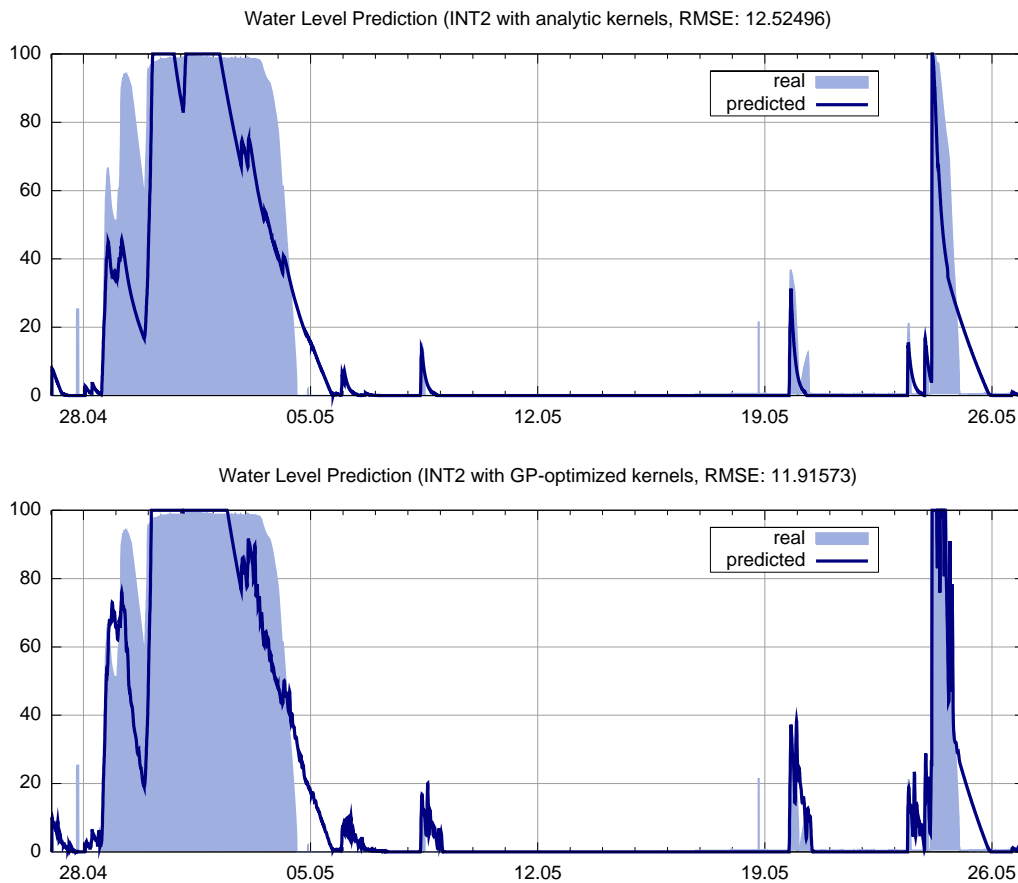


Abbildung 13: Visueller Vergleich von Vorhersagen mit analytischem und GP-optimiertem INT2 Modell.

- Finite Impulse Response Filter (FIR)
- Echo State Networks (ESN)
- Gewöhnliche Differentialgleichungen (ODE)
- Auf Integralgleichungen basierte Modelle (INT2)
- Support Vector Machines SVM
- Genetisches Programmieren GP
- Hybrid-Verfahren GP/INT2

Für einen unbefangenen Vergleich wurden alle Modellierungsmethoden zuerst mit SPO optimiert. Damit ist sicher gestellt, dass die Evaluation nicht durch die variierende Qualität der Default-Parametrisierung beeinflusst werden konnte. Unter den klassischen Ansätzen zeigte sich INT2 als am besten geeignet für die benötigten Vorhersagen, konnte aber durch das hybride GP/INT2 Verfahren sowie SVMs geschlagen werden. Als Beispiel für die Ergebnisse zeigt Abbildung 16 zum Vergleich die Vorhersagen eines mit SPO optimierten NARX Modells, und eines mit SPO optimierten GP/INT2 Modells.

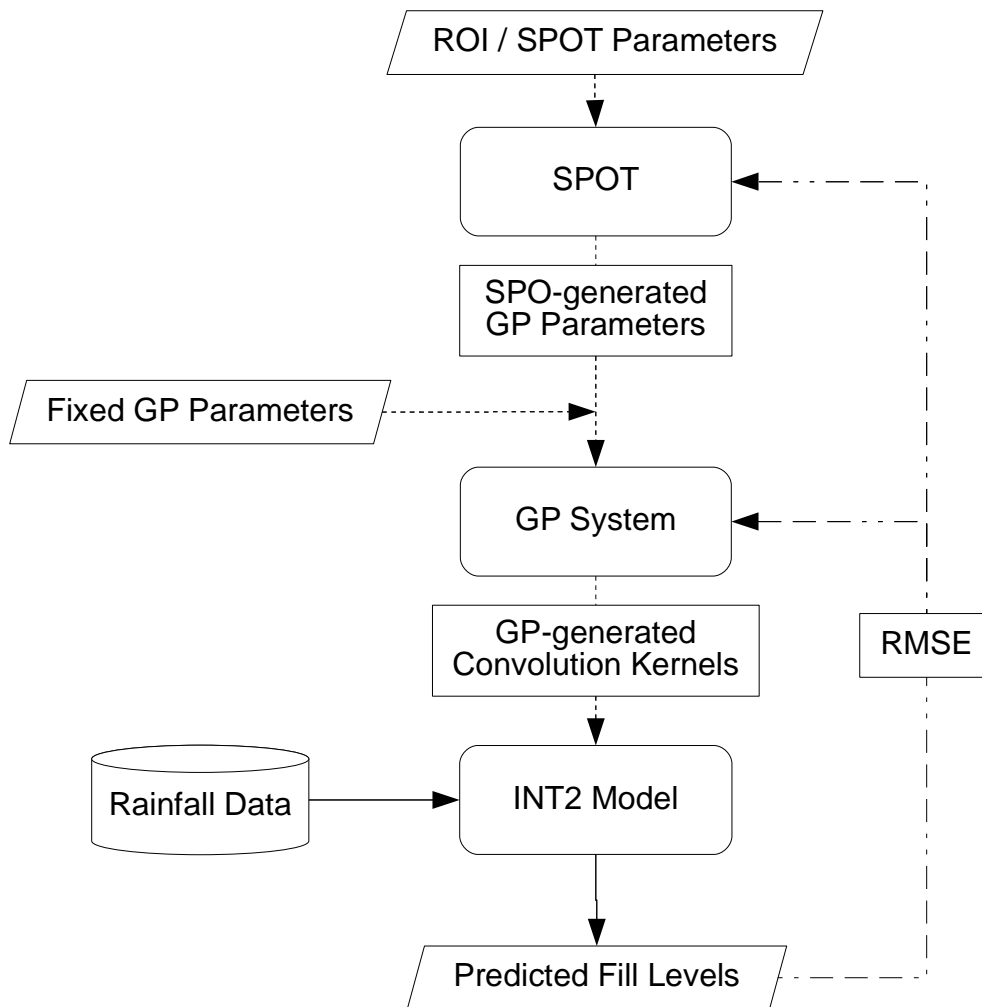


Abbildung 14: Schema des mit SPO optimierten hybriden GP/INT2 Ansatzes.

3.3.5 CS-3: Vergleichende Übersicht: Prognosemodelle in der Wasserwirtschaft

Diese Case Study wurde bereits im Berichtszeitraum 2009 im Rahmen des Masterstudiengangs “Automation & IT” von drei Studenten durchgeführt und erfolgreich abgeschlossen. Die Ergebnisse wurden in Form eines Konferenzbeitrags auf der Fachkonferenz „IEEE World Congress on Computational Intelligence (WCCI 2010)“ vorgestellt [30].

3.3.6 CS-4: Vergleich des FIWA-GP Prognosemodells mit anderen Modellen in der Wasserwirtschaft

Für diese Case Study konnten in der Projektlaufzeit leider keine geeigneten Studierenden gewonnen werden, so dass Arbeiten von dem im Projekt beschäftigten wissenschaftlichen Mitarbeiter Oliver Flasch übernommen werden mussten. Durch den Mehraufwand ergab sich eine Verzögerung beim Meilenstein MS-6 (Abnahme der Ergebnisse aus den Vergleichen). Die Vergleiche wurden im Rahmen des Arbeitspakets W-4 durchgeführt, siehe Abschnitt 3.3.4 für eine Übersicht der erzielten Ergebnisse.

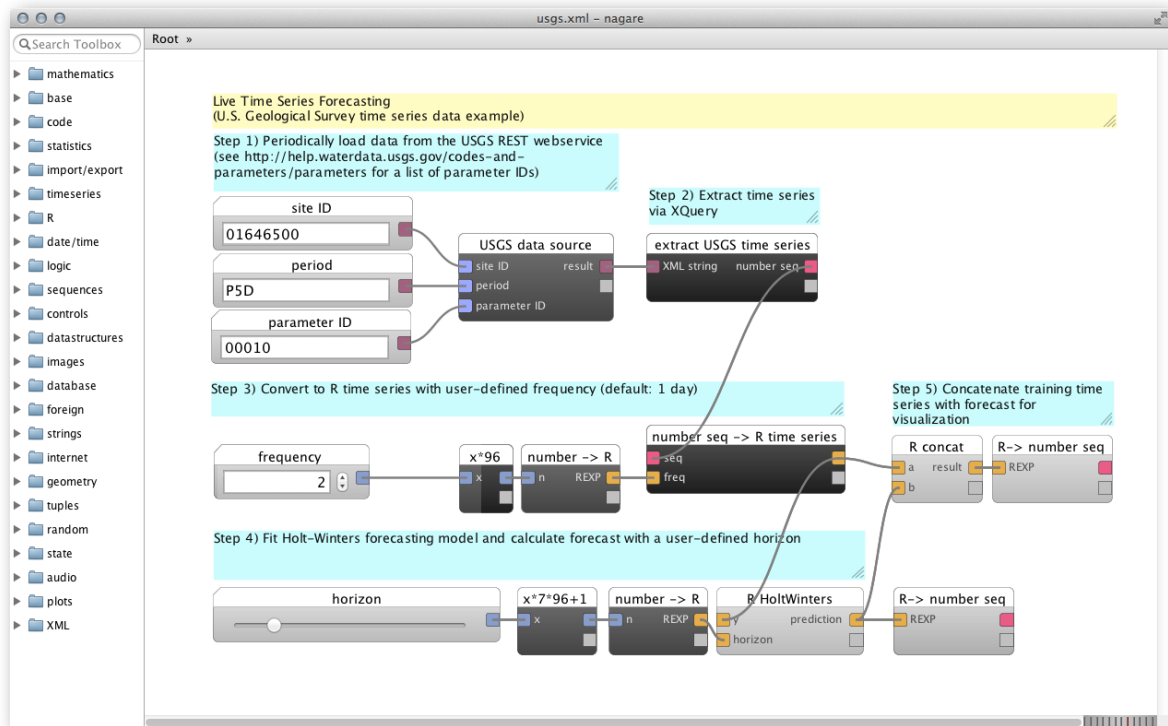


Abbildung 15: Werkzeug zur graphischen Darstellung, Manipulation und Auswertung von Prognosemodellen in der Wasserwirtschaft. Dargestellt ist ein System zur Echtzeit-Prognose von Flusspegelständen auf Basis von Zeitreihendaten des US Geographical Surveys (USGS). Das System empfängt Pegelstände in 15-Minuten-Intervallen, erstellt ein geeignetes Prognosemodell und wertet dieses Modell in Echtzeit aus.

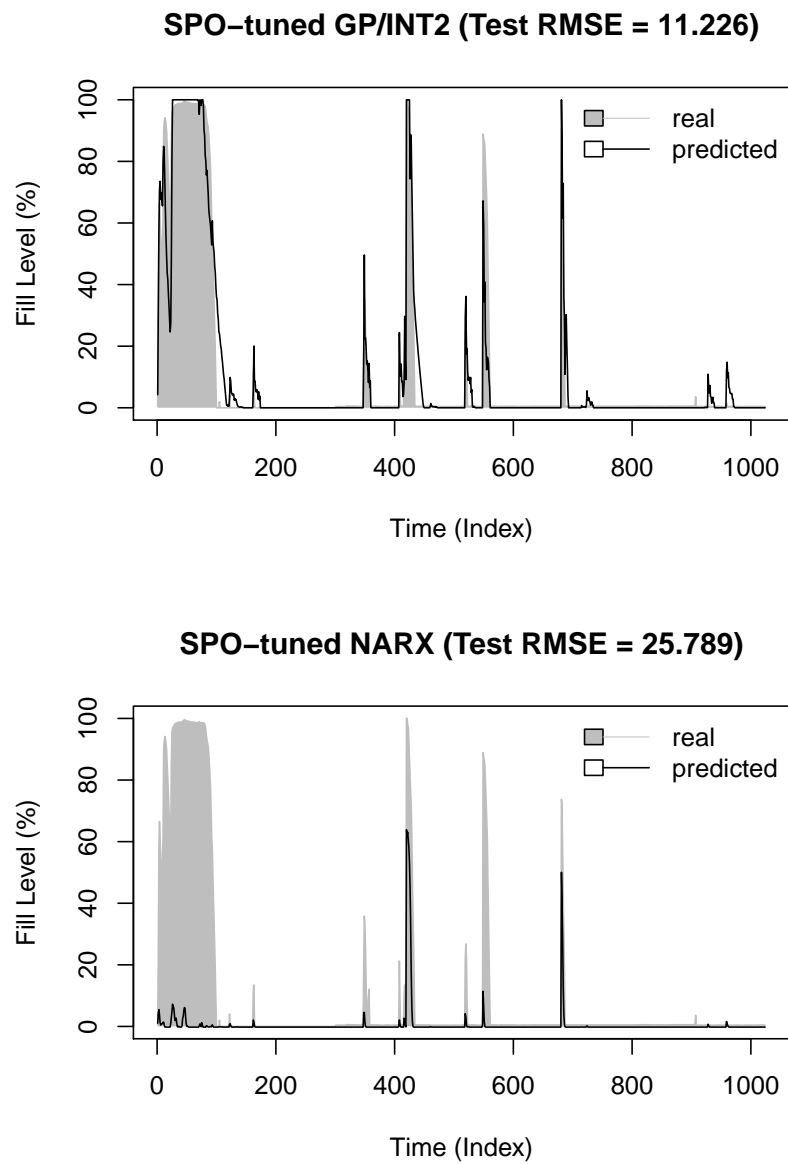


Abbildung 16: Visueller Vergleich von Vorhersagen mit SPO optimierten Modellen: NARX und GP/INT2

3.3.7 MT-2: Erstellung von Softwarepaketen für Wasserdaten

Da für die Bearbeitung dieser Masterarbeit keine geeigneten Studierenden gewonnen werden konnten, wurden die Aufgaben von dem im Projekt beschäftigten wissenschaftlichen Mitarbeiter übernommen. Durch den gesteigerten Arbeitsaufwand ergab sich eine Verzögerung im Projektabschluss. Die Inhalte dieser Masterarbeit wurden im Rahmen der Arbeitspakete W-2 und W-3 bearbeitet, siehe Abschnitte 3.3.2 und 3.3.3 für eine Beschreibung der erzielten Ergebnisse.

3.4 Modul M-4 „Generalisierbarkeit“

Ziel der Arbeitspakete im Modul M-4 war die Untersuchung der Generalisierbarkeit der auf Basis von Anwendungen aus der Finanz- und Wasserwirtschaft gefundenen Verfahren, Algorithmen und GP-System-Parametereinstellungen. Insbesondere erfolgte eine Einordnung der entwickelten Verfahren im Hinblick auf

- Zuverlässigkeit,
- Interpretierbarkeit,
- Robustheit,
- Anpassungsfähigkeit,
- Entwicklungs- und Wartungskosten.

Vergleichende experimentelle Analysen deckten allgemeine Eigenschaften von Zeitreihen aus der Finanz- und Wasserwirtschaft auf und konnten u.a. zur Entwicklung eines ersten allgemeinen Testproblem-Generators genutzt werden.

3.4.1 G-1: Unterschiede und Gemeinsamkeiten von GP-Systemen in der Finanz- und Wasserwirtschaft

In diesem Paket wurde untersucht, welche GP-Operatoren gleichermaßen für Problemstellungen der Finanz- und der Wasserwirtschaft geeignet sind und welche fundamentalen Unterschiede existieren. Der im Projekt beschäftigte wissenschaftliche Mitarbeiter war zur Konferenz *Genetic Programming in Theory and Practice (GPTP)*, welche im Mai 2012 Ann Arbor, Michigan (USA) stattfand, eingeladen. Die GPTP ist die renommierteste Konferenz im Bereich GP, eine Teilnahme ist nur auf Einladung möglich.

Auf dieser Konferenz wurde eine Vorgehensweise zur statistisch verifizierten Untersuchung von GP-Systemen vorgestellt [34]. Diese Vorgehensweise ermöglicht es in festgelegten Schritten, die sich an der Vorgehensweise der sequentiellen Parameteroptimierung SPO (vgl. Abschnitt 2.2) orientieren, Hypothesen zu überprüfen und statistisch abgesicherte Ergebnisse zu erzielen. In Bezug auf GP-Systeme war dazu der Entwurf von skalierbaren Testfunktionen notwendig, der ebenso in der Veröffentlichung vorgestellt wird. Die Generalisierbarkeit der Vorgehensweise wird durch die Untersuchung von zwei verschiedenen GP-Systemen belegt. Eine Übertragung auf das im Projekt entwickelte GP-System RGP (vgl. Abschnitt 3.1) fand ebenso bereits statt. Die Untersuchung weiterer Systeme ist ebenso möglich und sollte auf Basis der zur Verfügung gestellten Informationen ohne Probleme vollzogen werden können.

3.4.2 G-2: Generalisierbarkeit

Die in den Arbeitsgebieten Finanzzeitreihen und Wasserwirtschaft gefundenen Ergebnisse wurden zusammengetragen, um die Generalisierbarkeit des GP-Ansatzes zur Zeitreihenanalyse für unterschiedliche Problemstellungen zu beurteilen. Dabei wurde auch das Potenzial für weitere Anwendungsfelder (weitere Projektpartner, Fortsetzungsanträge) untersucht und neue Kooperationspartner gewonnen. Beispielsweise befindet sich ein Folgeprojekt mit dem Kooperationspartner GreenPocket GmbH in Vorbereitung. U.a. sollen in diesem Projekt Methoden aus FIWA auf die Prognose von Energieverbrauchs-Zeitreihen übertragen werden.

Eine erste Integration verschiedener Prognosemodelle, wie sie auch in FIWA entwickelt und verwendet wurden, bietet Flasch u. a. [38]. Hier werden verschiedene Modelle zu Ensembles zusammen gestellt mit denen dann Zeitreihen-Vorhersagen getätigt werden. Das besondere an diesen auf Ensembles basierten Verfahren ist die parallele Verwendung mehrerer Zeitreihenmodelle, die auf verschiedene Arten miteinander verknüpft oder auch einzeln verwendet werden können. Darüber hinaus kann eine Anpassung an neue Eigenschaften der betrachteten Zeitreihe automatisch erfolgen. In Flasch u. a. [38] wird ein auf Ensembles basiertes Verfahren mit einfachen Grundmodellen weiteren Verfahren gegenübergestellt und auf den Daten aus der GECCO INDUSTRIAL CHALLENGE (vgl. Abschnitt 2.5) verglichen.

Für die aufwändigen Berechnungen, die für die Erstellung von statistisch belastbaren Vergleichen nötig waren, kam ein Rechencluster der FH Köln zum Einsatz, dessen Hauptknoten (Head Node) und Rechenknoten (Compute Nodes) aus Mitteln des Projekts FIWA beschafft wurden. Die Installation und Konfiguration des Rechenclusters erfolgte durch Projektmitarbeiter. Das System wird auch weiterhin sehr aktiv für Forschung und Lehre genutzt. Seine Kapazität kann auf einfache Weise durch hinzufügen weiterer Rechenknoten erweitert werden.

Dieses Arbeitspaket diente auch der Vorbereitung der Dissertation,¹¹ für die ein eigenes Arbeitspaket vorgesehen war (vgl. hierzu Abschnitt 3.5). Das in diesem Projekt entstandene Material wurde gesichtet und rückblickend nach Relevanz geordnet, so dass die wesentlichen Elemente in die Dissertation einfließen können.

3.5 Spezielle Arbeitspakete

In diesem Abschnitt werden Ergebnisse und Stand der speziellen Arbeitspakete D (Dissertation) und K (Koordination) zusammengefasst.

D: Dissertation Die Erstellung der Dissertation, d.h. die abschließende Auswertung, Reflexion und Zusammenfassung, war als separates Modul eingeplant. Die Dissertation befindet sich in einem fortgeschrittenen Stadium. Einzelne Teile der Dissertation wurden erfolgreich zur Veröffentlichung bei Fachzeitschriften und Konferenzen eingereicht (siehe Abschnitt 3.6).

K: Koordination Während des Projektes fielen wichtige Koordinationsaufgaben an. Unter anderem wurden wöchentliche Statusseminare des Projektteams abgehalten. Dieses Arbeitspaket ist im Gantt-

¹¹Mit „Dissertation“ wird hier das eigentliche Werk (engl. thesis) bezeichnet.

Chart des Projektverlaufs separat ausgewiesen (siehe Abbildung 1).

3.6 Wissenschaftliche Publikationen

Durch die Mitarbeiter des Projektes FIWA wurden zahlreiche Publikationen veröffentlicht. Dabei zählt Prof. Dr. Bartz-Beielstein zu den Forschern der FH Köln mit den meisten begutachteten Publikationen. Im Projektzeitraum wurden drei Bücher [3, 15, 14], sowie zahlreiche Buchbeiträge [34, 27, 15, 21, 24, 27, 43], Seminar- und Abschlussarbeiten [1, 29, 56], Zeitschriftenaufsätze [47], Konferenzbeiträge [38, 18, 50, 46, 20, 42, 35, 4, 25, 30, 39, 31, 44, 45, 49, 57, 23, 37, 48], und Technische Berichte [33, 19, 2, 8, 9, 10, 11, 12, 13, 17, 26, 36, 58, 53, 5, 54] veröffentlicht. Mehrere eingeladene Vorträge und Tutorien [6, 7, 51, 16, 22, 32] und sonstige Vorträge [28, 52, 55] wurden im Projekt FIWA gehalten.

Die im Projektplan angekündigten und im Rahmen des Masterstudiengangs “Automation & IT” durchgeführten Case-Studies konnten erfolgreich abgeschlossen werden, wobei die Arbeiten für CS-4 von dem im Projekt beschäftigten wissenschaftlichen Mitarbeiter übernommen wurden. Des Weiteren wurde die Masterarbeit MT-1 erfolgreich fertiggestellt. Zusätzlich konnte eine weitere Masterarbeit im Projektrahmen initiiert und durchgeführt werden.

Im Rahmen der Konferenz *Genetic and Evolutionary Computation Conference* GECCO 2011 konnte der Wissenschaftswettbewerb GECCO INDUSTRIAL CHALLENGE zu Prognosemodellen in der Finanzwirtschaft mit großem Erfolg durchgeführt werden. Die Ergebnisse des Wettbewerbs wurden durch Flasch u. a. [35] beschrieben und analysiert. Auch im Folgejahr 2012 konnte die GECCO INDUSTRIAL CHALLENGE erneut erfolgreich durchgeführt werden [40] und ist auch für 2013 wieder geplant.

Die Nachwuchsförderung besitzt einen wichtigen Stellenwert im Projekt. Dazu gehört auch die regelmäßige Durchführung von Kolloquien.¹²

4 | Voraussichtlicher Nutzen

Das Projekt FIWA (Methoden der Computational Intelligence für Vorhersagemodelle in der Finanz- und Wasserwirtschaft) ist in einem sehr weit fortgeschrittenen Entwicklungsstadium, wichtige Meilensteine wurden plangemäß erreicht und interessante wissenschaftliche Ergebnisse konnten für Anwendungen sowohl der Finanzwirtschaft wie auch der Wasserwirtschaft erzeugt werden. Ein detaillierter Vergleich des Projektstands mit der ursprünglich veranschlagten Zeitplanung wurde bereits in Abschnitt 3 durchgeführt.

Durch die Nutzung der freien Statistiksoftware R konnte der Entwicklungsaufwand einiger Komponenten zur Zeitreihenanalyse signifikant gesenkt werden. Dies führte auch zu Ersparnissen bei der Beschaffung von Software. Durch den Einsatz von Open Source Software konnte die Reichweite der entwickelten Werkzeuge, auch für den Einsatz in der Lehre, stark erweitert werden. Gleichzeitig stieg jedoch durch den Wegfall des bei kommerzieller Software üblichen Supports der Arbeitsaufwand für den im Projekt beschäftigten wissenschaftlichen Mitarbeiter.

¹²Siehe http://fat70s236.eto.fh-koeln.de/ait/content/e183/e893/e1127/index_ger.html für eine Liste der Vorträge.

Die Arbeitspakete F-3, F-4, sowie W-3 und W-4 wurden bedingt durch die Ausrichtung der GECCO INDUSTRIAL CHALLENGE 2011 und 2012 mit etwas erhöhtem Aufwand bearbeitet.

Weiterer Nutzen lässt sich bereits in verschiedenen Bereichen beobachten. Hier sind zuerst der wissenschaftliche und der wirtschaftliche Bereich zu nennen. Die im Projekt FIWA entwickelten Methoden werden in verschiedenen Bereichen der Wirtschaft eingesetzt, eine konkrete Beschreibung der Methoden bei den jeweiligen Anwendern ist allerdings nicht möglich.

Darüber hinaus sind die neu entwickelten Methoden veröffentlicht worden, sodass sie auch in wissenschaftlicher Hinsicht Nutzen erzielt haben und weiter erzielen. Das FIWA GP-System, als typisiertes Graph-basiertes GP-System, ist für die Grundlagenforschung sehr interessant. Derartige GP-Systeme generieren Lösungen (Programme), die extern wie intern mit Daten verschiedener Typen (beispielsweise Zahlen, Zeichenreihen, Zeitreihen, etc.) arbeiten. Dies unterscheidet sie von untypisierten GP-Systemen und macht sie zudem für den Praxiseinsatz interessant. Basierend auf den entwickelten Methoden konnte das Interesse neuer Kooperationspartner geweckt werden, mit denen neue Projekte geplant oder bereits umgesetzt werden.

Das in der Vorhabenbeschreibung geplante FIWA-Nachwuchsteam ist etabliert worden, es setzt sich mittlerweile aus deutlich mehr Studenten und Promovierenden zusammen als damals geplant. Darüber hinaus finden Treffen nahezu wöchentlich statt. Mittlerweile ist das Nachwuchsteam in der Forschungsgruppe SPOTSEVEN aufgegangen, einen aktuellen Überblick über Themen und Aktivitäten liefert deren Web-Seite www.spotseven.de.

Die Inhalte und Ergebnisse des Projekts haben Einzug in verschiedene Vorlesungen genommen, so werden GP-Grundlagen mittlerweile in den Vorlesungen DDMO (*Data Driven Modeling and Optimization*) und APCO (*Advanced Process Control and Optimization*) des Antragstellers für den Master-Studiengang *Automation & IT* eingeführt und durch Beispiele aus dem Projekt ergänzt.

5 | Ergebnisse Dritter

Während der Durchführung des Vorhabens wurde auch bei anderen Stellen die Anwendung von GP in der Finanz- und erstmals auch in der Wasserwirtschaft erforscht. Mit einigen dieser externen Forschungsgruppen bestehen inzwischen Kooperationen.

Genetic Programming Das amerikanische Unternehmen Evolved Analytics LLC und sein belgischer Partner Evolved Analytics Europe BVBA entwickelten das kommerzielle GP-System DataModeler auf Basis des Computeralgebra-Systems *Mathematica* der Wolfram Research LLC. Kommerzielle Lizenzen für DataModeler sind in 2012 allgemein verfügbar.¹³ Das System wurde primär zur schnellen industriellen Datenanalyse, Dateninterpretation und Datenexploration multivariater Datensätze entworfen. DataModeler kann zur automatischen Erstellung von Zeitreihenprognosemodellen genutzt werden. Aus der Kooperation mit Evolved Analytics Europe sind gemeinsame Veröffentlichungen im Projektrahmen hervorgegangen. [38, 41]

Eine Ausgründung der amerikanischen Cornell University, die Nutonian Inc., entwickelte das kommerzielle GP-System Eurequa. Im Gegensatz zu DataModeler und RGP basiert das System nicht

¹³Siehe <http://www.evolved-analytics.com/?q=datamodeler> für Details.

auf einer bestehenden statistischen oder mathematischen Standardsoftware, sondern wird als „stand-alone“-Lösung genutzt. Eureka wurde primär zur Systemidentifikation entwickelt, d.h. zur Entdeckung von physikalischen Gesetzen auf Basis von multivariaten Messdaten.¹⁴ Mit Anpassungen kann das System auch für Anwendungen in der Finanz- und Wasserwirtschaft genutzt werden.

Finanzwirtschaft Im Bereich der Anwendung von GP in der Finanzwirtschaft sind drei Veröffentlichungen von Interesse. Wilson u. Banzhaf [99] untersuchen die Verwendung von Linear Genetic Programmig (LGP) für den automatischen Devisenhandel für vier Leitwährungen. Fitnessfunktionen mit verschiedener Realitätstreue, durch die optionale Berücksichtigung des höchsten Verlustes (Maximum Drawdown) wurden untersucht. Weiterhin wurden verschiedene Fitnessarten für verschiedene Marktsituationen untersucht im Hinblick ihrer Performanz über die Zeit, der erzeugten Handelsstrategien und ihrer Gesamtrentabilität. Die Analyse der Handelsrentabilität zeigt, dass das LGP-System in der Lage ist, mit moderater Handelsaktivität Profit zu erzielen und Verluste zu vermeiden.

In einer Folgeveröffentlichung untersuchen Wilson u. a. [100] die Verbesserung ihres LGP-Systems durch die Integration von Filtern. Unter anderem wird LGP auf verschiedene Intraday-Zeiträume angewandt, um ein Handelssignal zu erzeugen. Vier verschiedene Zeitreihen, die länger in die Vergangenheit reichen, werden untersucht. Zwei verschiedene Entscheidungsmechanismen für die Erzeugung des Gesamthandelssignals aus einem Ensemble von GP-Erzeugten Signalen aus allen Zeitbereichen werden ebenfalls untersucht, wobei einer auf Mehrheitsurteil und der andere auf der zeitgleichen Nähe des Kaufsignals beruht. Die Ergebnisse zeigen die Überlegenheit des Mehrheitsurteils. Weitere Analysen zeigen ausserdem, dass das Training auf längere Zeitreihen zu besseren Handelstrategien führt, sowohl in Märkten mit Aufwärtstrend wie auch in Märkten mit Abwärtstrend. [100]

Serpinis u. a. [95] vergleichen Psi Sigma Netze (PSI), eine moderne Variante künstlicher neuronaler Netze, mit dem Gene Expression algorithm (GEP), einer GP-Variante, bei der Vorhersage und dem Devisenhandel des EURUSD Wechselkurses. Als Benchmark dient die Performanz eines Multi-Layer Perceptron (MLP), einem rekurrenten künstlichen neuronalen Netzwerk, die Performanz eines einfachen GP Basisalgorithmus (GP), die Performanz eines ARMA-Modells, sowie die Performanz des naiven Modells, welches den letzten bekannten Wert als Vorhersage verwendet. Bei simulierten Handel erreichen PSI und GEP vergleichbare Performanz. Beide Modelle erzielen höhere Performanz als alle Benchmarks (MLP, GP, ARMA, naive), gemessen an jährlicher Rendite.

Wasserwirtschaft Im Bereich GP-Anwendungen in der Wasserwirtschaft ist eine Veröffentlichung von dritter Seite von Relevanz. Londhe u. Dixit [84] beschreiben ein System für die Echtzeit-Schätzung von Wellenhöhen im Golf von Mexiko auf GP-Basis. Das System wendet einen GP-Algorithmus an, um fehlende Daten zu Wellenhöhen an Messstationen anhand von Daten anderer Stationen in der Umgebung möglichst genau zu schätzen, es handelt sich also um ein multivariates Regressionsproblem. Im empirischen Teil wurden die Daten von sechs regionalen Messstation zur Schätzung der Daten von fünf benachbarten Messstationen herangezogen. Ein Vergleich der Güte der Schätzung mit anderen modernen CI-Methoden zeigt die Überlegenheit von GP bei diesem Anwendungsproblem.

¹⁴Siehe <http://www.nutonian.com/eureka/> für Details.

In dem ebenfalls in der Förderlinie Ingenieurwachstums 2009 geförderten Projekt SOMA (Prof. Dr. W. Konen, FH Köln) wurde erstmalig SPO auf klassische Methoden des Machine Learning angewandt. Durch intensive Zusammenarbeit mit diesem Projekt stand FIWA eine Anzahl moderner Prognosemodelle als Benchmark zur Verfügung. SOMA beteiligte sich an der Beschaffung von Rechenknoten des in Abschnitt 3.4.2 erwähnten Rechenclusters und nutzte diesen Cluster gemeinsam mit FIWA-Mitarbeitern. Dadurch konnte über die gesamte Projektlaufzeit und auch nach Projektende eine effiziente und hohe Nutzung der Ressourcen des Clusters sichergestellt werden.

6 | Veröffentlichungen im Projekt

- [1] AL-BADDAI, Saad ; AL-SUBARI, Karema: *Case-Study: Data Mining For Predictive Control In Environmental Technology*. February 2010
- [2] BARTZ-BEIELSTEIN, T: *Optimierung von Prozessvariablen beim Spritzgießen / Cologne University of Applied Sciences*. 2010. – Forschungsbericht
- [3] BARTZ-BEIELSTEIN, T. (Hrsg.) ; CHIARANDINI, M. (Hrsg.) ; PAQUETE, L. (Hrsg.) ; PREUSS, M. (Hrsg.): *Proceedings of Workshop on Experimental Methods for the Assessment of Computational Systems joint to PPSN2010*. 2010 (Computer Science series of the Dortmund University TR10-2-007)
- [4] BARTZ-BEIELSTEIN, Th. ; FLASCH, O. ; KOCH, P. ; KONEN, W.: SPOT: A Toolbox for Interactive and Automatic Tuning in the R Environment. In: HOFFMANN, F. (Hrsg.) ; HÜLLERMEIER, E. (Hrsg.): *Proceedings 20. Workshop Computational Intelligence*, Universitätsverlag Karlsruhe, 2010, S. 264–273
- [5] BARTZ-BEIELSTEIN, Thomas: Sequential Parameter Optimization. In: BRANKE, Jürgen (Hrsg.) ; NELSON, Barry L. (Hrsg.) ; POWELL, Warren B. (Hrsg.) ; SANTNER, Thomas J. (Hrsg.): *Sampling-based Optimization in the Presence of Uncertainty*. Dagstuhl, Germany : Schloss Dagstuhl - Leibniz-Zentrum fuer Informatik, Germany, 2009 (Dagstuhl Seminar Proceedings 09181). – ISSN 1862–4405
- [6] BARTZ-BEIELSTEIN, Thomas: *Automatic and Interactive Tuning of Algorithms—The Sequential Parameter Optimization Toolbox (Invited talk)*. Leiden University NL, December 2010
- [7] BARTZ-BEIELSTEIN, Thomas: *Experimental Analysis of Optimization Algorithms*. International Workshop on Nature Inspired Computation and Applications. University of Science and Technology of China, Oktober 2010
- [8] BARTZ-BEIELSTEIN, Thomas: Performing Experiments Using the Sequential Parameter Optimization Toolbox SPOT / Research Center CIOP (Computational Intelligence, Optimization and Data Mining). Version: July 2010. <http://maanvs03.gm.fh-koeln.de/webpub/CIOPReports.d/Bart10m.d/Bart10m.pdf>. Cologne University of Applied Science, Faculty of Computer Science and Engineering Science, July 2010 (06/10). – CIOP Technical Report. – ISSN 2191–365X
- [9] BARTZ-BEIELSTEIN, Thomas: Performing Meta Experiments Using the Sequential Parameter Optimization Toolbox SPOT / Cologne University of Applied Sciences. Version: Aug 2010. <http://maanvs03.gm.fh-koeln.de/webpub/CIOPReports.d/Bart10o.d/Bart10o.pdf>. Cologne University of Applied Science, Faculty of Computer Science and Engineering Science, Aug 2010 (08/10). – CIOP Technical Report. – ISSN 2191–365X

- [10] BARTZ-BEIELSTEIN, Thomas: Sequential Parameter Optimization—An Annotated Bibliography / Research Center CIOP (Computational Intelligence, Optimization and Data Mining). Version: April 2010. <http://maanvs03.gm.fh-koeln.de/webpub/CIOPReports.d/Bart10b.d/spotannotatedbib.pdf>. Cologne University of Applied Science, Faculty of Computer Science and Engineering Science, April 2010 (04/10). – CIOP Technical Report. – ISSN 2191–365X
- [11] BARTZ-BEIELSTEIN, Thomas: Severity Testing as a Meta-statistical Principle for Comparing Algorithms / Institute of Computer Science. Faculty of Computer Science and Engineering Science, Cologne University of Applied Sciences, Germany, 2010 (05/2010). – Forschungsbericht
- [12] BARTZ-BEIELSTEIN, Thomas: SPOT: An R Package For Automatic and Interactive Tuning of Optimization Algorithms by Sequential Parameter Optimization / Research Center CIOP (Computational Intelligence, Optimization and Data Mining). Version: June 2010. <http://arxiv.org/abs/1006.4645>. Cologne University of Applied Science, Faculty of Computer Science and Engineering Science, June 2010 (05/10). – CIOP Technical Report. – ISSN 2191–365X. – Comments: Related software can be downloaded from <http://cran.r-project.org/web/packages/SPOT/index.html>
- [13] BARTZ-BEIELSTEIN, Thomas: Writing Interfaces for the Sequential Parameter Optimization Toolbox SPOT / Cologne University of Applied Sciences. Version: July 2010. <http://maanvs03.gm.fh-koeln.de/webpub/CIOPReports.d/Bart10n.d/Bart10n.pdf>. Cologne University of Applied Science, Faculty of Computer Science and Engineering Science, July 2010 (07/10). – CIOP Technical Report. – ISSN 2191–365X
- [14] BARTZ-BEIELSTEIN, Thomas (Hrsg.) ; CHIARANDINI, Marco (Hrsg.) ; PAQUETE, Luis (Hrsg.) ; PREUSS, Mike (Hrsg.): *Experimental Methods for the Analysis of Optimization Algorithms*. Berlin, Heidelberg, New York : Springer, 2010. <http://dx.doi.org/10.1007/978-3-642-02538-9>. <http://dx.doi.org/10.1007/978-3-642-02538-9>
- [15] BARTZ-BEIELSTEIN, Thomas ; CHIARANDINI, Marco ; PAQUETE, Luis ; PREUSS, Mike: Introduction—Experimental Methods for the Analysis of Optimization Algorithms. In: BARTZ-BEIELSTEIN, Thomas (Hrsg.) ; CHIARANDINI, Marco (Hrsg.) ; PAQUETE, Luis (Hrsg.) ; PREUSS, Mike (Hrsg.): *Experimental Methods for the Analysis of Optimization Algorithms*. Berlin, Heidelberg, New York : Springer, 2010, S. 1–13
- [16] BARTZ-BEIELSTEIN, Thomas ; DAVIS, David ; MEHNEN, Jörn: *Evolutionary Computation in Practice track at Genetic and Evolutionary Computation Conference (GECCO)*. <http://www.isgec.org/gecco-2009/ecp.html>. Version: 2009
- [17] BARTZ-BEIELSTEIN, Thomas ; FLASCH, Oliver ; KOCH, Patrick ; KONEN, Wolfgang: SPOT: A Toolbox for Interactive and Automatic Tuning of Search Heuristics and Simulation Models in the R Environment / Cologne University of Applied Sciences. 2010. – Forschungsbericht
- [18] BARTZ-BEIELSTEIN, Thomas ; FLASCH, Oliver ; ZAEFFERER, Martin: Sequential Parameter Optimization for Symbolic Regression. In: GUSTAFSON, Steven (Hrsg.) ; VLADISLAVLEVA, Ekaterina (Hrsg.): *GECCO 2012 Symbolic regression and modeling workshop*. Philadelphia, Pennsylvania, USA : ACM, July 2012. – ISBN 978–1–4503–1178–6, S. 495–496
- [19] BARTZ-BEIELSTEIN, Thomas ; FRIESE, Martina ; FLASCH, Oliver ; KONEN, Wolfgang ; KOCH, Patrick ; NAUJOKS, Boris: Ensemble-Based Modeling / Research Center CIOP (Computational Intelligence, Optimization and Data Mining). Version: June 2011. <http://maanvs03.gm.fh-koeln.de/webpub/CIOPReports.d>

- d/Bart11e.d/Bart11e.pdf. Cologne University of Applied Science, Faculty of Computer Science and Engineering Science, June 2011 (06/11). – CIOP Technical Report. – ISSN 2191–365X
- [20] BARTZ-BEIELSTEIN, Thomas ; FRIESE, Martina ; ZAEFFERER, Martin ; NAUJOKS, Boris ; FLASCH, Oliver ; KONEN, Wolfgang ; KOCH, Patrick: Noisy optimization with sequential parameter optimization and optimal computational budget allocation. In: *Proceedings of the 13th annual conference companion on Genetic and evolutionary computation*. New York, NY, USA : ACM, 2011 (GECCO '11). – ISBN 978–1–4503–0690–4, 119–120
- [21] BARTZ-BEIELSTEIN, Thomas ; LASARCZYK, Christian ; PREUSS, Mike: The Sequential Parameter Optimization Toolbox. In: BARTZ-BEIELSTEIN, Thomas (Hrsg.) ; CHIARANDINI, Marco (Hrsg.) ; PAQUETE, Luis (Hrsg.) ; PREUSS, Mike (Hrsg.): *Experimental Methods for the Analysis of Optimization Algorithms*. Berlin, Heidelberg, New York : Springer, 2010, S. 337–360
- [22] BARTZ-BEIELSTEIN, Thomas ; PREUSS, Mike: *Experimental research in evolutionary computation–The Future of Experimental Research (Tutorial)*. Genetic and Evolutionary Computation Conf. (GECCO 2009), Montreal, CA, July, 2009
- [23] BARTZ-BEIELSTEIN, Thomas ; PREUSS, Mike: The future of experimental research. In: ROTHLAUF, Franz (Hrsg.): *GECCO (Companion)*, ACM, 2009. – ISBN 978–1–60558–505–5, S. 3185–3226
- [24] BARTZ-BEIELSTEIN, Thomas ; PREUSS, Mike: The Future of Experimental Research. In: BARTZ-BEIELSTEIN, Thomas (Hrsg.) ; CHIARANDINI, Marco (Hrsg.) ; PAQUETE, Luis (Hrsg.) ; PREUSS, Mike (Hrsg.): *Experimental Methods for the Analysis of Optimization Algorithms*. Berlin, Heidelberg, New York : Springer, 2010, S. 17–46
- [25] BARTZ-BEIELSTEIN, Thomas ; PREUSS, Mike: *Tuning and experimental analysis in evolutionary computation: what we still have wrong (Tutorial)*. Genetic and Evolutionary Computation Conf. (GECCO 2010), Portland, OR, July, 2010
- [26] BARTZ-BEIELSTEIN, Thomas ; PREUSS, Mike ; SCHMITT, Karlheinz ; SCHWEFEL, Hans-Paul: Challenges for Contemporary Evolutionary Algorithms / Faculty of Computer Science, Algorithm Engineering (Ls11), TU Dortmund. 2010 (TR10-2-003). – Technical Report
- [27] BARTZ-BEIELSTEIN, Thomas ; PREUSS, Mike ; SCHWEFEL, Hans-Paul: Model Optimization with Evolutionary Algorithms. In: LUCAS, K. (Hrsg.) ; ROOSEN, P. (Hrsg.): *Emergence, Analysis, and Evolution of Structures—Concepts and Strategies Across Disciplines*. Berlin, Heidelberg, New York : Springer, 2010, S. 47–62
- [28] DAVTYAN, Artur: *SPO-optimized NARX for fill-level prediction in stormwater tanks (Vortrag)*. COSA-Kolloquium, Cologne University of Applied Sciences, Februar 2010. – http://fat70s236.eto.fh-koeln.de/ait/content/e183/e893/e1127/index_ger.html
- [29] DAVTYAN, Artur ; OYETOYAN, Daniel T. ; TAMUTAN, Michael: *Case-Study: Fill-level prediction in stormwater tanks*. February 2010
- [30] FLASCH, O. ; BARTZ-BEIELSTEIN, Th. ; DAVTYAN, A. ; KOCH, P. ; KONEN, W. ; OYETOYAN, T.D. ; TAMUTAN, M.: Comparing SPO-tuned GP and NARX prediction models for stormwater tank fill level prediction. In: FOGEL, Gary et al. (Hrsg.): *Proc. IEEE Congress Evolutionary Computation (CEC)*, 2010, S. 1579–1586
- [31] FLASCH, O. ; BARTZ-BEIELSTEIN, Th. ; KOCH, P. ; KONEN, W.: Clustering Based Niching for Genetic Programming in the R Environment. In: HOFFMANN, F. (Hrsg.) ; HÜLLERMEIER, E. (Hrsg.): *Proceedings 20. Workshop Computational Intelligence*, Universitätsverlag Karlsruhe, 2010, S. 33–46

- [32] FLASCH, Oliver ; BARTZ-BEIELSTEIN, Thomas: *Sequential parameter optimization applied to evolutionary strategies for portfolio optimization*. Third International Conference on Computational and Financial Econometrics (CFE 09) and Second Workshop of the ERCIM Working Group on Computing & Statistics (ERCIM 09). Limassol, Cyprus, October 2009
- [33] FLASCH, Oliver ; BARTZ-BEIELSTEIN, Thomas: Towards a Framework for the Empirical Analysis of Genetic Programming System Performance / Research Center CIOP (Computational Intelligence, Optimization and Data Mining). Version: May 2012. <http://maanvs03.gm.fh-koeln.de/webpub/CIOPReports.d/Flas12a.d/ciop0512.pdf>. Faculty of Computer Science and Engineering Science, Cologne University of Applied Sciences, Germany, May 2012 (05/12). – CIOP Technical Report. – ISSN 2191–365X
- [34] FLASCH, Oliver ; BARTZ-BEIELSTEIN, Thomas: A Framework for the Empirical Analysis of Genetic Programming System Performance. In: RIOLO, Rick (Hrsg.) ; VLADISLAVLEVA, Ekaterina (Hrsg.) ; MOORE, Jason H. (Hrsg.): *Genetic Programming Theory and Practice X*. Ann Arbor, USA : Springer, Mai 2013 (in print) (Genetic and Evolutionary Computation)
- [35] FLASCH, Oliver ; BARTZ-BEIELSTEIN, Thomas ; 1, Daniel B. ; KANTSCHIK, Wolfgang ; STRACHWITZ, Christian von: Results of the GECCO 2011 Industrial Challenge: Optimizing Foreign Exchange Trading Strategies / Research Center CIOP (Computational Intelligence, Optimization and Data Mining). Version: December 2011. <http://maanvs03.gm.fh-koeln.de/webpub/CIOPReports.d/Flas11a.d/Flas11a.pdf>. Cologne University of Applied Science, Faculty of Computer Science and Engineering Science, December 2011 (10/11). – CIOP Technical Report. – ISSN 2191–365X
- [36] FLASCH, Oliver ; BARTZ-BEIELSTEIN, Thomas ; DAVTYAN, Artur ; KOCH, Patrick ; KONEN, Wolfgang ; OYE-TOYAN, Tosin D. ; TAMUTAN, Michael: Comparing CI Methods for Prediction Models in Environmental Engineering / Research Center CIOP (Computational Intelligence, Optimization and Data Mining). Version: Feb 2010. <http://maanvs03.gm.fh-koeln.de/webpub/CIOPReports.d/Flas10a.d/ciop0210.pdf>. Faculty of Computer Science and Engineering Science, Cologne University of Applied Sciences, Germany, Feb 2010 (02/10). – CIOP Technical Report. – ISSN 2191–365X
- [37] FLASCH, Oliver ; BARTZ-BEIELSTEIN, Thomas ; KOCH, Patrick ; KONEN, Wolfgang: Genetic Programming Applied to Predictive Control in Environmental Engineering. In: HOFFMANN, Frank (Hrsg.) ; HÜLLERMEIER, Eyke (Hrsg.): *Proceedings 19. Workshop Computational Intelligence*. Karlsruhe : KIT Scientific Publishing, 2009, S. 101–113
- [38] FLASCH, Oliver ; FRIESE, Martina ; VLADISLAVLEVA, Katya ; BARTZ-BEIELSTEIN, Thomas ; MERSMANN, Olaf ; NAUJOKS, Boris ; STORK, Joerg ; ZAEFFERER, Martin: Comparing Ensemble-based Forecasting Methods for Smart-Metering Data. In: *Conference on Applications of Evolutionary Computation (EvoApplications)*, Springer, 2013 (LNCS)
- [39] FLASCH, Oliver ; MERSMANN, Olaf ; BARTZ-BEIELSTEIN, Thomas: RGP: An Open Source Genetic Programming System for the R Environment. In: PELIKAN, Martin (Hrsg.) ; BRANKE, Jürgen (Hrsg.): *Genetic and Evolutionary Computation Conference, GECCO 2010, Proceedings, Portland, Oregon*, ACM, 2010, S. 2071–2072
- [40] FRIESE, M.: *GECCO 2012 Industrial Challenge*. <http://www.sigevo.org/gecco-2012/competitions.html#ic>. Version: 2012

- [41] FRIESE, Martina ; BARTZ-BEIELSTEIN, Thomas ; VLADISLAVLEVA, Katya ; FLASCH, Oliver ; MERSMANN, Olaf ; NAUJOKS, Boris ; ZAEFFERER, Martin ; STORK, Jörg: Ensemble-Based Model Selection for Smart Metering Data. In: HOFFMANN, Frank (Hrsg.) ; HÜLLERMEIER, Eyke (Hrsg.) ; Institut für Angewandte Informatik/Automatisierungstechnik am Karlsruher Institut für Technologie (Veranst.): *Proceedings 22. Workshop Computational Intelligence* Institut für Angewandte Informatik/Automatisierungstechnik am Karlsruher Institut für Technologie, KIT Scientific Publishing, 2012, S. 215–228
- [42] FRIESE, Martina ; ZAEFFERER, Martin ; BARTZ-BEIELSTEIN, Thomas ; FLASCH, Oliver ; KOCH, Patrick ; KONEN, Wolfgang ; NAUJOKS, Boris: Ensemble Based Optimization and Tuning Algorithms. In: HOFFMANN, F. (Hrsg.) ; HÜLLERMEIER, E. (Hrsg.): *Proceedings 21. Workshop Computational Intelligence*, Universitätsverlag Karlsruhe, 2011, S. 119–134
- [43] HUTTER, Frank ; BARTZ-BEIELSTEIN, Thomas ; HOOS, Holger ; LEYTON-BROWN, Kevin ; MURPHY, Kevin P.: Sequential Model-Based Parameter Optimisation: an Experimental Investigation of Automated and Interactive Approaches. In: BARTZ-BEIELSTEIN, Thomas (Hrsg.) ; CHIARANDINI, Marco (Hrsg.) ; PAQUETE, Luis (Hrsg.) ; PREUSS, Mike (Hrsg.): *Experimental Methods for the Analysis of Optimization Algorithms*. Berlin, Heidelberg, New York : Springer, 2010, S. 361–414
- [44] KOCH, Patrick ; KONEN, Wolfgang ; FLASCH, Oliver ; BARTZ-BEIELSTEIN, Thomas: Optimization of Support Vector Regression Models for Stormwater Prediction. In: HOFFMANN, F. (Hrsg.) ; HÜLLERMEIER, E. (Hrsg.): *Proceedings 20. Workshop Computational Intelligence*, Universitätsverlag Karlsruhe, 2010, S. 146–160
- [45] KOCH, Patrick ; KONEN, Wolfgang ; FLASCH, Oliver ; BARTZ-BEIELSTEIN, Thomas: Optimizing Support Vector Machines for Stormwater Prediction. In: BARTZ-BEIELSTEIN, T. (Hrsg.) ; CHIARANDINI, M. (Hrsg.) ; PAQUETE, L. (Hrsg.) ; PREUSS, M. (Hrsg.): *Proceedings of Workshop on Experimental Methods for the Assessment of Computational Systems joint to PPSN2010*. TU Dortmund, 2010 (TR10-2-007), 47–59
- [46] KOCH, Patrick ; KONEN, Wolfgang ; NAUJOKS, Boris ; FLASCH, Oliver ; FRIESE, Martina ; ZAEFFERER, Martin ; BARTZ-BEIELSTEIN, Thomas: Tuned Data Mining in R. In: HOFFMANN, F. (Hrsg.) ; HÜLLERMEIER, E. (Hrsg.): *Proceedings 21. Workshop Computational Intelligence*, Universitätsverlag Karlsruhe, 2011, S. 147–160
- [47] KONEN, W. ; ZIMMER, T. ; BARTZ-BEIELSTEIN, T.: Optimized Modelling of Fill Levels in Stormwater Tanks Using CI-based Parameter Selection Schemes (in german). In: *at-Automatisierungstechnik* 57 (2009), Nr. 3, S. 155–166. <http://dx.doi.org/10.1524/auto.2009.0756>. – DOI 10.1524/auto.2009.0756
- [48] KONEN, Wolfgang ; BARTZ-BEIELSTEIN, Thomas: Reinforcement learning for games: failures and successes. In: *GECCO '09: Proceedings of the 11th Annual Conference Companion on Genetic and Evolutionary Computation Conference*. New York, NY, USA : ACM, 2009, S. 2641–2648
- [49] KONEN, Wolfgang ; KOCH, Patrick ; FLASCH, Oliver ; BARTZ-BEIELSTEIN, Thomas: Parameter-Tuned Data Mining: A General Framework. In: HOFFMANN, F. (Hrsg.) ; HÜLLERMEIER, E. (Hrsg.): *Proceedings 20. Workshop Computational Intelligence*, Universitätsverlag Karlsruhe, 2010
- [50] KONEN, Wolfgang ; KOCH, Patrick ; FLASCH, Oliver ; BARTZ-BEIELSTEIN, Thomas ; FRIESE, Martina ; NAUJOKS, Boris: Tuned Data Mining: A Benchmark Study on Different Tuners. In: KRASNOGOR, Natalio (Hrsg.): *GECCO '11: Proceedings of the 13th Annual Conference on Genetic and Evolutionary Computation*, 2011, S. 1995–2002
- [51] MEHNEN, Jörn ; BARTZ-BEIELSTEIN, Thomas: *Evolutionary Computation in Practice track at Genetic and Evolutionary Computation Conference (GECCO)*. <http://www.sigevo.org/gecco-2010/ecp.html>. Version: 2010

- [52] OYETOYAN, Daniel T.: *Optimizing the GP/INT2 model for fill-level prediction in stormwater tanks with SPO (Vortrag)*. COSA-Kolloquium, Cologne University of Applied Sciences, Februar 2010. – http://fat70s236.eto.fh-koeln.de/ait/content/e183/e893/e1127/index_ger.html
- [53] STOEAN, Catalin ; PREUSS, Mike ; BARTZ-BEIELSTEIN, Thomas ; STOEAN, Ruxandra: *A New Clustering-Based Evolutionary Algorithm for Real-Valued Multimodal Optimization / Chair of Algorithm Engineering*. Version: 2009. http://ls11-www.cs.tu-dortmund.de/_media/techreports/tr09-07.pdf. TU Dortmund, Dortmund, Germany, 2009 (TR09-2-007). – Technical Report
- [54] STOEAN, Ruxandra ; BARTZ-BEIELSTEIN, Thomas ; PREUSS, Mike ; STOEAN, Catalin: *A Support Vector Machine-Inspired Evolutionary Approach for Parameter Setting in Metaheuristics / Research Center CIOP (Computational Intelligence, Optimization and Data Mining)*. Cologne University of Applied Science, Faculty of Computer Science and Engineering Science, Jan 2009 (01/09). – CIOP Technical Report. – ISSN 2191–365X
- [55] TAMUTAN, Michael: *GA-optimized NARX for fill-level prediction in stormwater tanks (Vortrag)*. COSA-Kolloquium, Cologne University of Applied Sciences, Februar 2010. – http://fat70s236.eto.fh-koeln.de/ait/content/e183/e893/e1127/index_ger.html
- [56] ZAEFFERER, Martin: *Simulation und Optimierung von Biogasanlagen mit Methoden der Computational Intelligence*, University of Applied Sciences Cologne, Faculty of Computer Science, Diplomarbeit, März 2010
- [57] ZIEGENHIRT, Jörg ; BARTZ-BEIELSTEIN, Thomas ; FLASCH, Oliver ; KONEN, Wolfgang ; ZAEFFERER, Martin: *Optimization of Biogas Production with Computational Intelligence—A Comparative Study*. In: FOGEL, Gary et a. (Hrsg.): *Proc. 2010 Congress on Evolutionary Computation (CEC'10) within IEEE World Congress on Computational Intelligence (WCCI'10), Barcelona, Spain*. Piscataway NJ : IEEE Press, 2010, S. 3606–3613
- [58] ZIEGENHIRT, Jörg ; BARTZ-BEIELSTEIN, Thomas ; FLASCH, Oliver ; KONEN, Wolfgang ; ZAEFFERER, Martin: *Optimization of Biogas Production with Computational Intelligence - A Comparative Study / Research Center CIOP (Computational Intelligence, Optimization and Data Mining)*. Version: Mar 2010. <http://maanvs03.gm.fh-koeln.de/webpub/CIOPReports.d/Zieg10a.d/ciop0310.pdf>. Cologne University of Applied Science, Faculty of Computer Science and Engineering Science, Mar 2010 (03/10). – CIOP Technical Report. – ISSN 2191–365X

7 | Literatur

- [59] BARANSKI, Bastian: *Evolutionsstrategien zur Portfoliooptimierung unter Nebenbedingungen*, Fachbereich Informatik, Universität Dortmund, Diplomarbeit, 2007
- [60] BARENDREGT, H. ; ABRAMSKY, S. ; GABBAY, D. M. ; MAIBAUM, T. S. E. ; BARENDREGT, H. P.: *Lambda Calculi with Types*. In: *Handbook of Logic in Computer Science*, Oxford University Press, 1992, S. 117–309
- [61] BARTZ-BEIELSTEIN, Thomas: *How Experimental Algorithmics Can Benefit from Mayo's Extensions to Neyman-Pearson Theory of Testing*. In: *Synthese* 163 (2008), Nr. 3, S. 385–396. – DOI10.1007/s11229-007-9297-z
- [62] BARTZ-BEIELSTEIN, Thomas: *Review: Design and Analysis of Simulation Experiments by Jack P.C. Kleijnen*. In: *INFORMS Computing Society News* 2 (2008), Fall 2008, S. 11–14

- [63] BARTZ-BEIELSTEIN, Thomas ; BONGARDS, Michael ; CLAES, Christoph ; KONEN, Wolfgang ; WESTENBERGER, Hartmut: Datenanalyse und Prozessoptimierung für Kanalnetze und Kläranlagen mit CI-Methoden. In: MIKUT, R. (Hrsg.) ; REISCHL, M. (Hrsg.): *Proc. 17th Workshop Computational Intelligence*, Universitätsverlag, Karlsruhe, 2007, S. 132–138
- [64] BARTZ-BEIELSTEIN, Thomas ; LASARCZYK, Christian ; PREUSS, Mike: Sequential Parameter Optimization. In: MCKAY, B. (Hrsg.) u. a.: *Proceedings 2005 Congress on Evolutionary Computation (CEC'05)*, Edinburgh, Scotland Bd. 1. Piscataway NJ : IEEE Press, 2005, S. 773–780
- [65] BARTZ-BEIELSTEIN, Thomas ; LASARCZYK, Christian ; PREUSS, Mike: Sequential Parameter Optimization Toolbox / Universität Dortmund, Germany. Version: 2008. <http://sfbc.uni-dortmund.de/Publications/Reference/Downloads/25608.pdf>. 2008. – Technical Report
- [66] BRABAZON, Anthony ; O'NEILL, Michael: *Biologically Inspired Algorithms for Financial Modelling*. Springer, 2006 (Natural Computing)
- [67] BROCKWELL, Peter J. ; DAVIS, Richard A.: *Introduction to Time Series and Forecasting*. New York NY : Springer, 2002
- [68] BURKHART, Martin: *Erprobung und Optimierung von Vorkläranlagenmodellen*, Universität Kaiserslautern, Diplomarbeit, 2002
- [69] CATTANEO, Giuseppe ; ITALIANO, Giuseppe: Algorithm engineering. In: *ACM Comput. Surv.* 31 (1999), Nr. 3, S. 3
- [70] COX, D. R.: *Planning of Experiments*. New York NY : Wiley, 1958
- [71] COX, D. R. ; HINKLEY, D. V.: *Theoretical Statistics*. London : Chapman and Hall, 1974
- [72] FUCHS, L. ; SEGELKE, K.: Gemeinsame Bewirtschaftung von Kanalnetz und Kläranlage. In: *ATV-DVWK-Workshop Mess- und Regelungstechnik in abwassertechnischen Anlagen*. Wuppertal, 2003, S. 1–9
- [73] GRÜNING, H.: Abflusssteuerung - quo vadis? In: *KA Korrespondenz Abwasser, Abfall* 55 (2008), S. 358–364
- [74] GRÜNING, H. ; ORTH, H. M.: Untersuchungen zur Effizienz einer Verbundsteuerung von Kanalnetz und Kläranlage auf der Basis innovativer Messtechniken. In: *KA-Abwasser, Abfall* (2004), Nr. 7
- [75] HACKING, I.: *Einführung in die Philosophie der Naturwissenschaften*. Stuttgart, Germany : Reclam, 1996
- [76] HILMER, T. ; BONGARDS, M.: Integrierte Steuer- und Regelstrategien für Kanalnetz und Kläranlage / Cologne University of Applied Sciences. 2008. – Forschungsbericht
- [77] HUI, A.: Using Genetic Programming to Perform Time-Series Forecasting of Stock Prices. In: KOZA, John R. (Hrsg.): *Genetic Algorithms and Genetic Programming at Stanford 2003*. Stanford, CA : Stanford Bookstore, 2003, S. 83–90
- [78] IFAK: *Simulationssystem SIMBA*. <http://www.ifak-system.de/produkte/simulationssoftware/abwassersimulation/simba-6.html>, 2008
- [79] KANTSCHIK, Wolfgang: *Genetische Programmierung und Schach*, Universität Dortmund, Diss., 2006

- [80] KOZA, J.R.: *Genetic Programming: On the Programming of Computers by Means of Natural Selection*. Cambridge MA : MIT Press, 1992
- [81] LASARCZYK, Christian W. G.: *Genetische Programmierung einer algorithmischen Chemie*, Technische Universität Dortmund, Diss., 2007
- [82] LASARCZYK, Christian W. G. ; BANZHAF, Wolfgang: An Algorithmic Chemistry for Genetic Programming. In: KEIJZER, Maarten (Hrsg.) u. a.: *Proceedings of the 8th European Conference on Genetic Programming* Bd. 3447. Berlin, Heidelberg, New York : Springer, 2005 (Lecture Notes in Computer Science), S. 1–12
- [83] LASARCZYK, Christian W. G. ; BANZHAF, Wolfgang: Total synthesis of algorithmic chemistries. In: *GECCO '05: Proceedings of the 2005 Annual Conference on Genetic and Evolutionary Computation*. New York, NY, USA : ACM, 2005. – ISBN 1–59593–010–8, S. 1635–1640
- [84] LONDHE, Shreenivas N. ; DIXIT, Pradnya R.: Genetic Programming: A Novel Computing Approach in Modeling Water Flows. Version: 2012. <http://dx.doi.org/doi:10.5772/48179>. In: VENTURA, Sebastian (Hrsg.): *Genetic Programming - New Approaches and Successful Applications*. InTech, 2012. – DOI doi:10.5772/48179, Kapitel 9, S. 199–224
- [85] MAYO, D. G.: *Error and the Growth of Experimental Knowledge*. Chicago IL : The University of Chicago Press, 1996
- [86] MEHNEN, Jörn ; MICHELITSCH, Thomas ; LASARCZYK, Christian ; BARTZ-BEIELSTEIN, Thomas: Multi-objective evolutionary design of mold temperature control using DACE for parameter optimization. In: *International Journal of Applied Electromagnetics and Mechanics* 25 (2007), Nr. 1–4, 661–667. <http://iospress.metapress.com/content/751K5GG10P79Q501>
- [87] MEHNEN, Jörn ; MICHELITSCH, Thomas ; LASARCZYK, Christian W. G. ; BARTZ-BEIELSTEIN, Thomas: Multi-objective Evolutionary Design of Mold Temperature Control using DACE for Parameter Optimization. In: PFÜTZNER, H. (Hrsg.) ; LEISS, E. (Hrsg.): *Proceedings Twelfth International Symposium Interdisciplinary Electromagnetics, Mechanics, and Biomedical Problems (ISEM 2005)* Bd. L11-1. Vienna, Austria : Vienna Magnetics Group Reports, 2005, S. 464–465
- [88] PARKER, Wendy: Computer simulation through an error-statistical lens. In: *Synthese* 163 (2008), Nr. 3, S. 371–384
- [89] POLI, Riccardo ; LANGDON, William B. ; MCPHEE, Nicholas F.: *A field guide to genetic programming*. Published via <http://lulu.com> and freely available at <http://www.gp-field-guide.org.uk>, 2008. – (With contributions by J. R. Koza)
- [90] POPPER, K.: *The Logic of Scientific Discovery*. London, U.K. : Hutchinson, 1959
- [91] SANTNER, T. J. ; WILLIAMS, B. J. ; NOTZ, W. I.: *The Design and Analysis of Computer Experiments*. Berlin, Heidelberg, New York : Springer, 2003
- [92] SCHILLING, W. (Hrsg.): *Praktische Aspekte der Abflusssteuerung in Kanalnetzen*. München : Oldenbourg, 1996
- [93] SCHLITTEGEN, Rainer: *Zeitreihenanalyse*. Oldenbourg, 2001

- [94] SCHMIDT, Michael ; LIPSON, Hod: Comparison of tree and graph encodings as function of problem complexity. In: THIERENS, Dirk (Hrsg.) ; BEYER, Hans-Georg (Hrsg.) ; BONGARD, Josh (Hrsg.) ; BRANKE, Jurgen (Hrsg.) ; CLARK, John A. (Hrsg.) ; CLIFF, Dave (Hrsg.) ; CONGDON, Clare B. (Hrsg.) ; DEB, Kalyanmoy (Hrsg.) ; DOERR, Benjamin (Hrsg.) ; KOVACS, Tim (Hrsg.) ; KUMAR, Sanjeev (Hrsg.) ; MILLER, Julian F. (Hrsg.) ; MOORE, Jason (Hrsg.) ; NEUMANN, Frank (Hrsg.) ; PELIKAN, Martin (Hrsg.) ; POLI, Riccardo (Hrsg.) ; SASTRY, Kumara (Hrsg.) ; STANLEY, Kenneth O. (Hrsg.) ; STUTZLE, Thomas (Hrsg.) ; WATSON, Richard A. (Hrsg.) ; WEGENER, Ingo (Hrsg.): *GECCO '07: Proceedings of the 9th Annual Conference on Genetic and Evolutionary Computation* Bd. 2. London : ACM Press, 7-11 July 2007, S. 1674–1679
- [95] SERMPINIS, Georgios ; LAWS, Jason ; KARATHANASOPOULOS, Andreas ; DUNIS, Christian L.: Forecasting and trading the EUR/USD exchange rate with Gene Expression and Psi Sigma Neural Networks. In: *Expert Systems with Applications* 39 (2012), Nr. 10, S. 8865–8877. <http://dx.doi.org/doi:10.1016/j.eswa.2012.02.022>. – DOI doi:10.1016/j.eswa.2012.02.022. – ISSN 0957–4174
- [96] SMITS, G. ; VLADISLAVLEVA, E.: Ordinal Pareto Genetic Programming. In: YEN, Gary G. (Hrsg.) u. a.: *Proceedings of the 2006 IEEE Congress on Evolutionary Computation*. Vancouver, BC, Canada : IEEE Press, 16-21 Juli 2006, 3114–3120
- [97] TSAY, R. S.: *Analysis of Financial Time Series*. New York, NY : Wiley, 2005
- [98] U.S. ENVIRONMENTAL PROTECTION AGENCY: *Storm Water Management Model*. 2007. – <http://www.epa.gov/ednrmr1/models/swmm/index.htm>, Online; Stand 26.08.07
- [99] WILSON, Garnett ; BANZHAF, Wolfgang: Interday foreign exchange trading using linear genetic programming. In: AL., Juergen B. (Hrsg.): *GECCO '10: Proceedings of the 12th annual conference on Genetic and evolutionary computation*. Portland, Oregon, USA : ACM, 7-11 Juli 2010, S. 1139–1146
- [100] WILSON, Garnett ; LEBLANC, Derek ; BANZHAF, Wolfgang: Stock trading using linear genetic programming with multiple time frames. In: AL., Natalio K. (Hrsg.): *GECCO '11: Proceedings of the 13th annual conference on Genetic and evolutionary computation*. Dublin, Ireland : ACM, 12-16 Juli 2011, S. 1667–1674
- [101] YAN, W. ; CLACK, C. D.: Evolving robust GP solutions for hedge fund stock selection in emerging markets. In: THIERENS, D. (Hrsg.) ; BEYER, H.-G. (Hrsg.) ; BONGARD, J. (Hrsg.) u. a.: *GECCO '07: Proc. 9th Conf. Genetic and Evolutionary Computation*. New York, NY : ACM Press, 2007, S. 2234–2241

Abkürzungsverzeichnis

ACM	Association for Computing Machinery
AR	Auto Regressive
ARMA	Auto Regressive Moving Average
CI	Computational Intelligence
CIOP	Forschungsstelle für Computational Intelligence, Optimization & Data Mining
E.R.R.O.R	Experimental Reasoning, Reliability, Objectivity and Rationality
EA	Evolutionäre Algorithmen
EC	Evolutionary Computation

ESN	Echo State Networks
FIR	Finite Impulse Response
GECCO	Genetic and Evolutionary Computation Conference
GP	Genetic Programming
KMU	Kleine und Mittlere Unternehmen
MA	Moving Average
MOSR	Multi-Objective Symbolic Regression
NARX	Nonlinear AutoRegressive neural networks with eXogenous inputs
NN	Neuronale Netze
ODE	Ordinary Differential Equations
RMSE	Root Mean Squared Error
SE	Squared Error
SPO	Sequentielle Parameteroptimierung
SPOT	Sequential Parameter Optimization Toolbox
STEPS	Sustainable Technologies and Computational Services for Environmental and Production Processes
SVM	Support Vector Machine
USGS	United States Geographical Survey

GEFÖRDERT VOM



Bundesministerium
für Bildung
und Forschung

Die Verantwortung für den Inhalt dieser
Veröffentlichung liegt beim Autor.

Kontakt/Impressum

Diese Veröffentlichungen erscheinen im Rahmen der Schriftenreihe "Ciplus". Alle Veröffentlichungen dieser Reihe können unter

www.ciplus-research.de

oder unter

<http://opus.bsz-bw.de/fhk/index.php?la=de>

abgerufen werden.

Köln, Januar 2012

Herausgeber / Editorship

Prof. Dr. Thomas Bartz-Beielstein,

Prof. Dr. Wolfgang Konen,

Prof. Dr. Horst Stenzel,

Dr. Boris Naujoks

Institute of Computer Science,

Faculty of Computer Science and Engineering Science,

Cologne University of Applied Sciences,

Steinmüllerallee 1,

51643 Gummersbach

url: www.ciplus-research.de

Schriftleitung und Ansprechpartner/ Contact editor's office

Prof. Dr. Thomas Bartz-Beielstein,

Institute of Computer Science,

Faculty of Computer Science and Engineering Science,

Cologne University of Applied Sciences,

Steinmüllerallee 1, 51643 Gummersbach

phone: +49 2261 8196 6391

url: <http://www.gm.fh-koeln.de/~bartz/>

eMail: thomas.bartz-beielstein@fh-koeln.de

ISSN (online) 2194-2870