

**Genetisches Programmieren als neues
Instrumentarium zur Prognose
makroökonomischer Größen: Anwendungen auf
Inflationsraten und Wechselkurse**

Inaugural-Dissertation zur Erlangung des Grades eines Doktors der Wirtschafts-
und Sozialwissenschaften der Wirtschafts- und Sozialwissenschaftlichen Fakultät
der Christian-Albrechts-Universität zu Kiel

vorgelegt von
Diplom-Volkswirt Elmar Zschischang

Kiel, Januar 2005

Gedruckt mit Genehmigung der
Wirtschafts- und Sozialwissenschaftlichen Fakultät
der Christian-Albrechts-Universität zu Kiel

Dekan: Professor Dr. Andreas Drexl
Referent: Professor Dr. Thomas Lux
Korreferent: Professor Dr. Gerd Hansen

Tag des Abschlusses der Arbeit: 24. Januar 2005
Tag der mündlichen Prüfung: 27. April 2005

Inhaltsverzeichnis

Tabellenverzeichnis	VII
Abbildungsverzeichnis	XII
Einleitung	1
1 Erfolge nichtlinearer Modelle auf Finanzmärkten	4
1.1 Maschinelles Lernen	6
1.2 Künstliche Neuronale Netze	7
1.3 Fuzzy System	9
1.4 Evolutionary Computation	10
1.4.1 Evolutionsstrategien	10
1.4.2 Evolutionary Programming	11
1.4.3 Genetische Algorithmen	13
1.5 Anwendungen computerbasierter Suchalgorithmen auf den Finanzmärkten .	14
1.6 Zusammenfassung	17
2 Genetisches Programmieren	18
2.1 Aufbau und Ablauf des Genetischen Programms	19
2.2 Vorteile des Genetischen Programms	23
2.3 Schema Theorem	24
2.3.1 GP Schema Theorem nach Poli & Langdon	25
2.4 Bloating	30
2.5 Residuenanalyse	31
2.5.1 Test auf Normalverteilung	31
2.5.2 Test auf Autokorrelation	31
2.6 Symbolische Regressionen und Vergleich des GP GAUSS mit anderen Pro- grammen	32
2.7 Sensitivitätsanalyse	37
2.7.1 Variation der Funktionsmenge	38
2.7.2 Unterschiedliche Gewichtung genetischer Operationen	40
2.7.3 Variation der Populationsgröße und der Anzahl der Generationen . .	42

2.7.4	Variation der Länge der Gleichungen in der Anfangspopulation . . .	43
2.7.5	Variation der maximal erlaubten Länge der Gleichungen	44
2.7.6	Selektionsmethode für genetische Operationen	44
2.7.7	Verwenden verschiedener Fitneßkriterien	48
2.7.8	Ansatz zur Analyse der Building Blocks	49
2.7.9	Implementieren eines mit Backpropagation bezeichneten Optimierungsverfahrens	54
2.7.10	Zusammenfassung	55
3	Geldpolitik und die Einordnung nichtlinearer Schätzmethoden	57
3.1	Ausgestaltung der Geldpolitik	57
3.2	Inflation Targeting	59
3.3	Genetisch entwickelte Modelle - ein geldpolitisches Prognoseinstrument?	60
3.4	Methoden zur Gewinnung von Inflationserwartungen	60
3.4.1	Umfragen	61
3.4.2	Finanzmarktdaten	61
4	Inflationsprognosemodelle	63
4.1	Mishkin Ansatz	63
4.2	Ergebnisse ausgewählter empirischer Studien	65
4.3	Konstruktion von Vergleichsmodellen auf Basis deutscher Monatsdaten . . .	68
4.3.1	Test auf Stationarität	68
4.3.2	Tests auf die Prognosegüte	69
4.3.3	Random Walk	71
4.3.4	Modellierung von Long Memory Effekten: ARIMA vs. ARFIMA . .	71
4.3.5	VECM	78
4.3.6	Out-of-Sample Prognosen anhand der konstruierten Vektorfehler- korrekturmodelle	84
4.4	Genetisch entwickelte Prognosemodelle	87
4.4.1	In-Sample und Out-of-Sample	89
4.4.2	In-Sample, Trainingsphase und Out-of-Sample	96
4.4.3	In-Sample, Trainingsphase, Out-of-Sample plus Backpropagation . .	96
4.4.4	In-Sample, Trainingsphase, Out-of-Sample und Building Blocks . . .	98
4.4.5	In-Sample, Trainingsphase, Out-of-Sample und Informationskriterium	102
4.4.6	Prognose von Inflationsänderungen	103
4.5	Zusammenfassung	105

5 Wechselkursprognosemodelle	108
5.1 Theoretische Ansätze zur Bestimmung des Wechselkursverhaltens	109
5.1.1 Kaufkraftparitätentheorie	109
5.1.2 Zinsparitätentheorie	110
5.1.3 Monetäre Modelle der Wechselkursentwicklung	111
5.1.4 Monetäre Theorie bei flexiblen Preisen	112
5.1.5 Monetäre Theorie bei kurzfristig trägen Preisen	113
5.1.6 Portfoliotheorie	114
5.2 Daten	115
5.3 Zeitreiheneigenschaften	115
5.4 Lang- und kurzfristige rekursiv geschätzte Modelle	116
5.5 GP GAUSS	117
5.6 Ergebnisse der lang- und kurzfristigen Prognosemodelle	119
5.7 Zusammenfassung	122
Zusammenfassung und Ausblick	128
A Tests auf Nichtlinearität	130
A.1 McLeod & Li	130
A.2 Engle	130
A.3 Tsay	131
A.4 Hinich Bikovarianz Test	131
A.5 Hinich Bispektraltest	132
A.6 BDS Test	132
B Symbolische Regression	137
B.1 Ergebnisse deterministisch dynamischer Prozesse	137
B.1.1 Lineare Ozaki Funktion	137
B.1.2 Logistic Map	139
B.1.3 Henon Map	140
B.1.4 Nichtlineare Funktion von Tong	142
B.1.5 Mackey-Glass	143
B.2 Ergebnisse deterministisch stochastischer Prozesse	145
B.2.1 Bilineares System von Granger & Andersen	145
B.2.2 Grenzwert Modell TAR von Liu	147
B.3 Ergebnisse bei Schätzen von Pseudo-Zufallszahlen	148
C Weitere Auswertungen der Inflationsprognosemodelle	151
C.1 Schätzung von Zinsstrukturkurven - Svensson Methode	151

C.2	ARIMA	152
C.3	ARFIMA	153
C.4	Genetische Programme: weitere Ergebnisse zu Inflationsprognosen	154
D	Weitere Auswertungen der Wechselkursmodelle	160
D.1	Daten für Wechselkursmodelle	160
D.2	GP GAUSS: langfristige Wechselkursprognosen	166
D.2.1	MSE Fitneßkriterium: Punkt- und Richtungsprognosen	166
D.2.2	AIC Fitneßkriterium: Punkt- und Richtungsprognosen	168
D.3	GP GAUSS: kurzfristige Wechselkursprognosen	170
D.3.1	AIC vs. MSE - Punktprognose	170
D.3.2	AIC vs. MSE - Richtungsprognose	171
D.4	Langfristige Wechselkursprognosen: Auszug aus der Studie von Verkooijen et al.	176
	Literaturverzeichnis	177
	Danksagung	177

Tabellenverzeichnis

2.1	Anfangsparameter des GP GAUSS Grundmodells	22
2.2	Vergleich der Ergebnisse von GP GAUSS mit denen von Kaboudan & Vance	33
2.3	Vergleich der Ergebnisse von GP GAUSS mit denen von Kaboudan	36
2.4	Sensitivitätsanalyse: Variation der Funktionsmengen	38
2.5	Sensitivitätsanalyse: Variation der Populationsgröße	42
2.6	Sensitivitätsanalyse: Variation der Anfangspopulation	44
2.7	Sensitivitätsanalyse: Variation der Selektionsmethode	48
2.8	Sensitivitätsanalyse: Building Blocks - Vergleich Crossover und Mutation .	53
4.1	Inflationsmodelle: Test auf Stationarität	69
4.2	Richtungsprognose: Aufbau der Kontingenztabelle	71
4.3	Random Walk Prognosen	72
4.4	Inflationsmodelle: ARIMA(2,1,3) und ARIMA(2,1,0)	73
4.5	Inflationsmodelle: Gedächtnisparameter für Inflationsraten	76
4.6	Inflationsmodelle: Gedächtnisparameter für differenzierte Inflationsraten . .	77
4.7	Inflationsmodelle: Gedächtnisparameter	77
4.8	Anpassung der Modellordnung anhand des BIC Kriteriums	78
4.9	Parameterschätzungen eines ARFIMA(3, d , 1) Modells	78
4.10	Prognosen verschiedener ARFIMA Spezifikationen	79
4.11	Auswahl der Laglänge für das VECM	81
4.12	VECM: Bestimmung der Eigenwerte	82
4.13	Stationaritätstest der Kointegrationsvektoren	83
4.14	Companion Matrix	84
4.15	Anpassungsparameter der Kointegrationsvektoren	85
4.16	VECM: Prognosewerte bei Konstanz der durchschnittlichen Inflationsrate .	86
4.17	Zusammenfassung der Prognosen der Vergleichsmodelle	87
4.18	Anfangsparameter des GP GAUSS Grundmodells	89
4.19	GP GAUSS: In-Sample und Out-of-Sample	90
4.20	MSE Entwicklung der größten prozentualen Verbesserungen aus Abbildung	90
4.21	Besten ex-post Prognosen	94

4.22	Richtungsprognosen der Vergleichsmodelle	95
4.23	In-Sample, Trainingsphase, Out-of-Sample: besten Punktprognosen	96
4.24	In-Sample, Trainingsphase, Out-of-Sample: besten Richtungsprognosen	97
4.25	Backpropagation: besten Punktprognosen	97
4.26	Backpropagation: besten Richtungsprognosen	98
4.27	Building Block Analyse	100
4.28	Building Block	101
4.29	Informationskriterium	102
4.30	Informationskriterium: Richtungsprognosen	102
4.31	Prognose von Inflationsänderungen	105
5.1	Test auf Stationarität	115
5.2	Anfangsparameter des GP GAUSS Grundmodells	118
5.3	Langfristige Wechselkurspunktprognosen anhand AIC ab 1974	119
5.4	Langfristige Wechselkurspunktprognosen anhand MSE ab 1974	120
5.5	Langfristige Wechselkursrichtungsprognosen anhand MSE ab 1974	124
5.6	Langfristige Wechselkursrichtungsprognosen anhand AIC ab 1974	125
5.7	Kurzfristige Wechselkursrichtungsprognosen ab 1974	126
5.8	Kurzfristige Wechselkurspunktprognosen ab 1974	127
A.1	BDS Testergebnisse nach Kanzler	134
A.2	Verschiedene nichtlineare Testergebnisse: Deutschland und Niederlande	135
A.3	Verschiedene nichtlineare Testergebnisse: Japan, Großbritannien und U.S.A.	136
B.1	OZ Residuenanalyse	138
B.2	LG Residuenanalyse	140
B.3	HN Residuenanalyse	142
B.4	SH Residuenanalyse	142
B.5	MG Residuenanalyse	145
B.6	BL Residuenanalyse	146
B.7	TAR Residuenanalyse	149
B.8	UR Residuenanalyse	149
C.1	Inflationsprognose: Modellspezifikation anhand AIC	152
C.2	Inflationsprognose: Modellspezifikation anhand SIC	152
C.3	Inflationsprognose: Analyse der Schätzresiduen für ARIMA(2,1,0)	153
C.4	Inflationsprognose: Analyse der Schätzresiduen für ARIMA(2,1,3)	153
C.5	Parameterschätzungen eines ARFIMA(1, 0.3799, 1) Modells	153
C.6	Parameterschätzungen eines ARFIMA(1, d , 1) Modells	153

C.7	2.Simulationslauf auf der Basis eines MSE Fitneßkriteriums: verschiedene Auswahlstrategien	154
C.8	2.Simulationslauf auf der Basis eines AIC Fitneßkriteriums: verschiedene Auswahlstrategien	154
C.9	Auswahlkriterium: besten In-Sample MSE Werte	156
C.10	Auswahlkriterium: die besten MSE Trainingwerte	157
C.11	Auswahlkriterium: Building Block < 1%	158
C.12	Auswahlkriterium: Building Block < 1% - Trainingsrichtungsprognosen . . .	159
D.1	Wechselkursmodelle: verwendete Zeitreihen - Datenquelle	165
D.2	Langfristige Wechselkurspunktprognosen anhand MSE ab 1987	166
D.3	Langfristige Wechselkursrichtungsprognosen anhand MSE ab 1987	167
D.4	Langfristige Wechselkurspunktprognosen anhand AIC ab 1987	168
D.5	Langfristige Wechselkursrichtungsprognosen anhand AIC ab 1987	169
D.6	Kurzfristige Wechselkurspunktprognosen ab 1987	170
D.7	Kurzfristige Wechselkursrichtungsprognosen ab 1987	171
D.8	Prognoseergebnisse von Verkooijen et al.	176

Abbildungsverzeichnis

1.1	Forschungszweige des Maschinellen Lernens	5
1.2	Aufbau eines künstlichen Neuronales Netzwerkes	8
1.3	Finite State Machine	12
1.4	Evolutionsoperatoren bei Genetischen Algorithmen	14
2.1	Genetisches Programmieren: Crossover	19
2.2	Genetisches Programmieren: Mutation	20
2.3	Flußdiagramm für Genetisches Programmieren	21
2.4	Verschiedene Schemata	24
2.5	Schema Theorem: Schema nach Poli & Langdon	25
2.6	Schema Theorem: Zerstörung eines Schemas	28
2.7	Schätzung der trigonometrischen Funktion TF mit GP GAUSS	35
2.8	Sensitivitätsanalyse: Variation der Funktionsmengen - Balkendiagramm . .	39
2.9	Sensitivitätsanalyse: Variation der Funktionsmengen - Kurvendiagramm . .	39
2.10	Sensitivitätsanalyse: Variation Crossoveranteil - Balkendiagramm	41
2.11	Sensitivitätsanalyse: Variation Crossoveranteil - Fitneßentwicklung	41
2.12	Sensitivitätsanalyse: Variation der Populationsgröße	43
2.13	Sensitivitätsanalyse: Variation der Anfangspopulation - Knotenentwicklung	45
2.14	Sensitivitätsanalyse: Variation der Anfangspopulation - Fitneßentwicklung .	45
2.15	Sensitivitätsanalyse: Variation der Knotenlänge	46
2.16	Sensitivitätsanalyse: Variation der Selektionsmethode	47
2.17	Sensitivitätsanalyse: Alternative Selektionskriterien	49
2.18	Building Block Analyse: Generation 1 bis 1000	51
	(a) MSE Fitneßentwicklung	51
	(b) Fitneßänderungen: Crossoveroperationen im Zeitablauf	51
	(c) Fitneßänderungen: Mutationsoperationen im Zeitablauf	51
	(d) Fitneßänderungen ausgetauschter Teilstränge in %	51
2.19	Building Block Analyse: Generation 450 bis 700	52
	(a) MSE Fitneßentwicklung	52
	(b) Fitneßänderungen: Crossoveroperationen im Zeitablauf	52

(c)	Fitneßänderungen: Mutationsoperationen im Zeitablauf	52
(d)	Fitneßänderungen ausgetauschter Teilstränge in %	52
2.20	Sensitivitätsanalyse: Backpropagation	54
(a)	MSE Fitneßentwicklung	54
2.21	Sensitivitätsanalyse: Backpropagation - Fitneßverbesserung in %	55
3.1	Zins- und Inflationsentwicklung von Januar 1973 bis Dezember 2000	62
4.1	Out-of-Sample Prognosen der Vergleichsmodelle	88
(a)	Inflationsprognose - ARIMA(2,d,0)	88
(b)	Inflationsprognose - ARFIMA(1,d,1)	88
(c)	Inflationsprognose - VECM RLZ 2 Jahre	88
(d)	Inflationsprognose - Random Walk	88
4.2	Prozentuale Verbesserung zwischen Generation 200 und 1000	91
4.3	MSE Prognosewerte zwischen Generation 1 und 200	92
4.4	MSE Prognosewerte zwischen Generation 1 und 1000	92
4.5	Tabelle 4.19: Lauf 66, Generation 1000	93
4.6	Tabelle 4.19: Lauf 4, Generation 200	93
4.7	Tabelle 4.21: Lauf 21, Generation 1000 (Modell B)	93
4.8	Tabelle 4.21: Lauf 66, Generation 200 (Modell A)	93
4.9	Prognosegüte eines ausgewählten Laufs	95
4.10	Building Block Analyse	99
(a)	In-Sample: Fitneßentwicklung	99
(b)	Out-of-Sample: Prognoseentwicklung	99
(c)	In-Sample: Generation 1 bis 50	99
(d)	Fitneßänderungen durch Crossover	99
(e)	Out-of-Sample: Generation 1 bis 50	99
(f)	Fitneßänderungen durch Crossoveroperationen: Generation 1 bis 50	99
4.11	Building Block Analyse: Histogramm der Punktprognosen	101
4.12	Informationskriterium: Histogramm der Punktprognosen	103
4.13	Prognose von Inflationsänderungsraten	104
(a)	Informationskriterium: Histogramm der Inflationsänderungsraten	104
(b)	MSE: Histogramm der Inflationsänderungsraten	104
B.1	Dynamischer Prozeß: OZ - beste GP GAUSS Regression	137
B.2	Histogramm OZ - die besten MSE Werte aus 100 Läufen	138
B.3	Dynamischer Prozeß: LG - beste GP GAUSS Regression	139
B.4	Histogramm LG - die besten MSE Werte aus 100 Läufen	140
B.5	Dynamischer Prozeß: HN - beste GP GAUSS Regression	141

B.6	Histogramm HN - die besten MSE Werte aus 100 Läufen	141
B.7	Dynamischer Prozeß: SH - beste GP GAUSS Regression	143
B.8	Histogramm SH - die besten MSE Werte aus 100 Läufen	143
B.9	Dynamischer Prozeß: MG - beste GP GAUSS Regression	144
B.10	Histogramm MG - die besten MSE Werte aus 100 Läufen	145
B.11	Dynamischer Prozeß: BL - beste GP GAUSS Regression	146
B.12	Histogramm BL - die besten MSE Werte aus 100 Läufen	147
B.13	Dynamischer Prozeß: TAR - beste GP GAUSS Regression	148
B.14	Histogramm TAR - die besten MSE Werte aus 100 Läufen	148
B.15	Dynamischer Prozeß: UR - beste GP GAUSS Regression	150
B.16	Histogramm UR - die besten MSE Werte aus 100 Läufen	150
C.1	Histogramme Inflationsprognosen: Abschnitt 4.4.2 und 4.4.3	155
	(a) Inflationsprognose: In-Sample-, Trainings- und Out-of-Sample Menge .	155
	(b) Inflationsprognose: Backpropagation Modul	155
C.2	Histogramme Inflationsprognosen: Wiederholte GP GAUSS Simulation . . .	155
	(a) 2.Simulationslauf: Inflationsprognose auf der Basis eines MSE Fitneß-	
	kriteriums; Pop = 100, Gen = 200, Runs = 100	155
	(b) 2.Simulationslauf: Inflationsprognose auf der Basis eines AIC Fitneß-	
	kriteriums; Pop = 100, Gen = 200, Runs = 100	155
D.1	Zeitreihen zur Schätzung von Wechselkursmodellen: Deutschland	161
	(a) Logarithmierter Wechselkurs	161
	(b) Inflation	161
	(c) Zinsdifferenz	161
	(d) Logarithmierte Geldmengendifferenz	161
	(e) Differenz der logarithmierten Industrie Produktion	161
	(f) Handelsbilanz BRD	161
D.2	Zeitreihen zur Schätzung von Wechselkursmodellen: Japan	162
	(a) Logarithmierter Wechselkurs	162
	(b) Inflation	162
	(c) Zinsdifferenz	162
	(d) Logarithmierte Geldmengendifferenz	162
	(e) Differenz der logarithmierten Industrie Produktion	162
	(f) Handelsbilanz Japan	162
D.3	Zeitreihen zur Schätzung von Wechselkursmodellen: Großbritannien	163
	(a) Logarithmierter Wechselkurs	163
	(b) Inflation	163
	(c) Zinsdifferenz	163

(d)	Logarithmierte Geldmengendifferenz	163
(e)	Differenz der logarithmierten Industrie Produktion	163
(f)	Handelsbilanz GB	163
D.4	Zeitreihen zur Schätzung von Wechselkursmodellen: Niederlande	164
(a)	Logarithmierter Wechselkurs	164
(b)	Inflation	164
(c)	Zinsdifferenz	164
(d)	Logarithmierte Geldmengendifferenz	164
(e)	Differenz der logarithmierten Industrie Produktion	164
(f)	Handelsbilanz NL	164
D.5	Wechselkursvorhersagen für DM/US- $\text{\$}$: Prognosehorizont 24 Monaten . . .	172
(a)	Monetäre Theorie bei flexiblen Preisen: DM/US- $\text{\$}$ Prognose mit MSE als Fitneßkriterium	172
(b)	Monetäre Theorie bei flexiblen Preisen: DM/US- $\text{\$}$ Prognose mit AIC als Fitneßkriterium	172
(c)	Monetäre Theorie bei trägen Preisen: DM/US- $\text{\$}$ Prognose mit MSE als Fitneßkriterium	172
(d)	Monetäre Theorie bei trägen Preisen: DM/US- $\text{\$}$ Prognose mit AIC als Fitneßkriterium	172
(e)	Portfoliotheorie: DM/US- $\text{\$}$ Prognose mit MSE als Fitneßkriterium . .	172
(f)	Portfoliotheorie: DM/US- $\text{\$}$ Prognose mit AIC als Fitneßkriterium . .	172
D.6	Wechselkursvorhersagen für Yen/US- $\text{\$}$: Prognosehorizont 24 Monaten . . .	173
(a)	Monetäre Theorie bei flexiblen Preisen: Yen/US- $\text{\$}$ Prognose mit MSE als Fitneßkriterium	173
(b)	Monetäre Theorie bei flexiblen Preisen: Yen/US- $\text{\$}$ Prognose mit MSE als Fitneßkriterium	173
(c)	Monetäre Theorie bei trägen Preisen: Yen/US- $\text{\$}$ Prognose mit MSE als Fitneßkriterium	173
(d)	Monetäre Theorie bei trägen Preisen: Yen/US- $\text{\$}$ Prognose mit AIC als Fitneßkriterium	173
(e)	Portfoliotheorie: Yen/US- $\text{\$}$ Prognose mit MSE als Fitneßkriterium . .	173
(f)	Portfoliotheorie: Yen/US- $\text{\$}$ Prognose mit AIC als Fitneßkriterium . .	173
D.7	Wechselkursvorhersagen für Pfund/US- $\text{\$}$: Prognosehorizont 24 Monaten . .	174
(a)	Monetäre Theorie bei flexiblen Preisen: Pfund/US- $\text{\$}$ Prognose mit MSE als Fitneßkriterium	174
(b)	Monetäre Theorie bei flexiblen Preisen: Pfund/US- $\text{\$}$ Prognose mit AIC als Fitneßkriterium	174
(c)	Monetäre Theorie bei trägen Preisen: Pfund/US- $\text{\$}$ Prognose mit MSE als Fitneßkriterium	174

(d)	Monetäre Theorie bei trägen Preisen: Pfund/US-\$ Prognose mit AIC als Fitneßkriterium	174
(e)	Portfoliotheorie: Pfund/US-\$ Prognose mit MSE als Fitneßkriterium .	174
(f)	Portfoliotheorie: Pfund/US-\$ Prognose mit AIC als Fitneßkriterium .	174
D.8	Wechselkursvorhersagen für GFL/US-\$: Prognosehorizont 24 Monaten . . .	175
(a)	Monetäre Theorie bei flexiblen Preisen: GFL/US-\$ Prognose mit MSE als Fitneßkriterium	175
(b)	Monetäre Theorie bei flexiblen Preisen: GFL/US-\$ Prognose mit AIC als Fitneßkriterium	175
(c)	Monetäre Theorie bei trägen Preisen: GFL/US-\$ Prognose mit MSE als Fitneßkriterium	175
(d)	Monetäre Theorie bei trägen Preisen: GFL/US-\$ Prognose mit AIC als Fitneßkriterium	175
(e)	Portfoliotheorie: GFL/US-\$ Prognose mit MSE als Fitneßkriterium . .	175
(f)	Portfoliotheorie: GFL/US-\$ Prognose mit AIC als Fitneßkriterium .	175

Einleitung

Seit den 1970er Jahren haben die internationalen Finanzmärkte erheblich an Bedeutung gewonnen. Die Gründe für das starke Wachstum sind vielfältig. Für die Devisenmärkte ist das Ende des Bretton-Woods Wechselkurssystems im Jahr 1973 ein einschneidendes Ereignis gewesen. Durch den Übergang zu einem flexiblen Dollarkurs sind größere Wechselkurs- und Zinsschwankungen entstanden, aufgrund derer international tätige Unternehmen in wachsendem Maße Finanzgeschäfte zur Absicherung dieser Schwankungsrisiken abschlossen. Die zahlreichen Finanzinnovationen in den 1980er Jahren, die Beseitigung von Beschränkungen im Kapitalverkehr, der zunehmende Interbankenhandel, die Einführung technischer Neuheiten - wie computergestützte Handelssysteme oder das Internet als zusätzliche Plattform für Handel und Informationsaustausch - haben wesentlich zur Ausweitung des Handelsvolumens beigetragen. Trotz des stark veränderten Umfeldes sind die Gruppierungen der Finanzakteure gleichgeblieben: Händler, Spekulanten, Privatanleger, institutionelle Anleger und Zentralbanken. Die Intentionen, die hinter den Transaktionen der einzelnen Gruppen stehen, sowie ihr Anlagehorizont unterscheiden sich teilweise sehr deutlich voneinander. So besteht die Aufgabe der Händler zwar primär in der Aufrechterhaltung eines liquiden Marktes, aber sie verfolgen ebenso das Ziel, kurzfristige Ineffizienzen des Marktes in Form von Arbitragegewinnen auszunutzen. Zusätzlich obliegt ihnen teilweise die Aufgabe, im Eigenhandel mit Hilfe technischer Indikatoren mögliche Kurstrends ausfindig zu machen, um Gewinne zu erzielen. Der Anlagehorizont dieser spekulativen Transaktionen beträgt lediglich wenige Minuten. Daher benötigen die Intra-Tagesgeschäfte möglichst gute Prognosemodelle, die die Risiken einzuschränken versuchen. Der Anlagehorizont der Spekulanten und Privatanleger hängt von deren Präferenzen ab, wobei der Spekulant eher kurzfristig und der Privatanleger eher mittel- bis langfristig orientiert ist. Die institutionellen Anleger und die Zentralbanken agieren in der Regel langfristig. Institutionelle Anleger wie die international tätigen Unternehmen sichern sich beispielsweise auf dem Devisenmarkt gegen Wechselkursrisiken ab, die erheblichen Einfluß auf deren Gewinn- und Umsatzentwicklung ausüben können. In diesem Zusammenhang läßt sich auch das Interesse der Zentralbanken an einem stabilen Außenwert ihrer Währung begründen. Der Außenwert, der die Kaufkraft der inländischen Währung im Ausland ausdrückt, spiegelt sich in der Wechselkursentwicklung wider. Somit sind sowohl institutionelle Anleger als auch die Zentralbanken an der langfristigen Entwicklung der Wechselkurse interessiert. Gute Prognosemodelle könnten dabei Grundlage für zukünftige wirtschaftspolitische oder unternehmerische Entscheidungen sein. Die Hauptaufgabe der Zentralbanken besteht allerdings in der Schaffung von Preis(niveau)-

stabilität¹. Dieses Ziel ist bei vielen Zentralbanken entwickelter Volkswirtschaften gesetzlich verankert. Für das Erreichen stabiler Preise bilden die Finanzmärkte die wichtigste Grundlage für die Geldpolitik. Die Regulierung der Zinsen und der Geldmenge dienen als Instrumente der Zentralbanken. Ihre Entscheidungen basieren dabei auf theoretisch entwickelten Modellen wie dem monetaristischen oder keynesianisch-neoklassischen Ansatz oder auf finanzmarktbasierenden Inflationserwartungen, wie sie sich beispielsweise in der Zinsstrukturkurve wiederfinden. Die Zentralbanken verfolgten in den letzten Jahrzehnten nicht immer die gleiche Strategie für diese Aufgabe. Seit Anfang der 1990er Jahre sind einige Zentralbanken zu einer neuen geldpolitischen Konzeption des sogenannten „Inflation Targeting“ übergegangen², das als ein zentrales Element Inflationsprognosen erstellt. Die geldpolitische Ausgestaltung der europäischen Zentralbank beruht dagegen auf einer Zwei-Säulen-Strategie, anhand derer sie das primäre Ziel der Preisstabilität verfolgt. Die Grundlage der ersten Säule bilden monetäre Analysen mit der zentralen Aussage, daß die Inflation auf lange Sicht stets eine monetäre Erscheinung ist. Die zweite Säule fußt auf Konjunktur- und Finanzindikatoren, die sich auf reale ökonomische Variablen stützen³. Während die ökonometrischen Modelle in Bezug auf Inflationsraten zumeist bessere Vorhersagen erzeugen als das naive Modell eines Random Walks⁴, gilt dies im Hinblick auf Wechselkursprognosen nicht⁵. Dabei besagt der Random Walk nichts weiter, als daß der beste Prognosewert unabhängig vom Prognosehorizont der aktuelle Wert einer Zeitreihe ist.

Die vorliegende Arbeit möchte dazu beitragen, Prognosemodelle auf der Grundlage des Genetischen Programmierens zu erstellen, in der Hoffnung, die Entscheidungsfindung zu erleichtern. Dieser junge Forschungsbereich imitiert die Methoden des natürlichen Evolutionsprozesses und ist in der Lage, selbständig den unbekanntem Datengenerierungsprozeß einer beliebigen Zeitreihe in Form einer mathematischen Funktion zu approximieren. Anschließend werden anhand der gefundenen Funktion Vorhersagen getroffen und deren Güte mit konventionellen ökonometrischen Modellen verglichen. Motiviert wird die vorliegende Arbeit vor allem durch die zahlreichen Erfolge auf dem Gebiet der Evolutionary Computation, zu dem das Genetische Programmieren gehört.

Die Arbeit ist in vier große Abschnitte eingeteilt. Im ersten Abschnitt erfolgt neben der Einordnung des Genetischen Programmierens in den Gesamtzusammenhang des maschinellen Lernens die Darstellung einiger erfolgreicher Anwendungen auf Finanzmarktdaten. Sie geben lediglich einen kleinen Ausschnitt der Anwendungsmöglichkeiten computerbasierter Optimierungsalgorithmen wieder. Gute bibliographische Zusammenfassungen der verschiedenen Anwendungsmöglichkeiten insbesondere im Bereich der Ökonomie und der Finanzmärkte geben die Bücher von Chen [21], [20] und Ribeiro et al. [109]. Im zweiten Abschnitt wird das mit dem Matrizenrechenprogramm GAUSS eigenständig programmierte Genetische Computerprogramm einem Gütevergleich mit anderen frei verfügbaren Gene-

¹Die Europäische Zentralbank beispielsweise versteht hierunter, daß die jährliche Wachstumsrate des harmonisierten Verbraucherpreisindex unterhalb von 2% und mittelfristig unterhalb aber nahe 2% betragen soll. Monthly Bulletin Mai 2003 der EZB: <http://www.ecb.int/pub/pdf/mobu/mb200305en.pdf>.

²Siehe auch Bofinger et al. [11], S. 365 ff.

³Vgl. auch Monthly Bulletin November 2000, Europäische Zentralbank.

⁴Vgl. auch Blix et al. [10].

⁵Vgl. auch Meese et al. [89], Cheung et al [23] oder Plasmans et al. [104].

tischen Programmen unterzogen. Eine daran anschließende Sensitivitätsanalyse soll dabei helfen, das Genetische Programm zu optimieren. Im dritten Abschnitt wird das Genetische Programm auf ein bekanntes theoretisches Modell zur Inflationsprognose angewandt, um verschiedene Strategien für die ex-ante Auswahl der Inflationsprognosegleichung aus einer Menge an potentiellen Lösungskandidaten zu überprüfen. Dabei handelt es sich um den Ansatz von Mishkin, der den Informationsgehalt von nominalen Zinsdifferenzen im Hinblick auf zukünftige Änderungen der Inflationsrate untersucht. Der vierte Abschnitt setzt sich mit Wechselkursprognosen auseinander. Grundlage bilden die Studien von Meese & Rogoff [89], Plasmans et al. [104] und Verkooijen et al. [127], die die klassischen und nach wie vor populären Wechselkursmodelle der 1970er und 1980er Jahre auf ihre Prognosequalität hin analysiert haben. Während Meese und Rogoff hierzu konventionelle lineare Regressionsmethoden herangezogen haben, benutzten Plasmans et al. und Verkooijen et al. [127] künstliche neuronale Netzwerke zur Schätzung nichtlinearer Prognosemodelle. Sämtlichen Untersuchungen liegen monatliche Beobachtungswerte und somit niedrig frequentierte Zeitreihen zugrunde.

Kapitel 1

Erfolge nichtlinearer Modelle auf Finanzmärkten

Das Interesse an nichtlinearen Regressionsmethoden zur Erstellung von Finanzmarktmodellen ist in den letzten Jahren sprunghaft angewachsen. Meese & Rogoff haben mit ihrer bekannten Studie im Jahr 1983, in der sie lineare Strukturmodelle zur Wechselkursvorhersage angewendet haben, zu dieser Entwicklung beigetragen. Sie untersuchten die in den 1970er Jahren konzipierten monetaristischen, keynesianischen und portfoliobasierten Wechselkursmodelle sowie weitere uni- und multivariate Zeitreihenmodelle mit dem Ergebnis, bei Prognosehorizonten von eins, sechs und zwölf Monaten keine signifikant besseren Vorhersagen zu erzeugen als das naive Modell eines Random Walks. Diese Ergebnisse wurden in der Literatur häufig als Bestätigung der schwachen Effizienzmarkthypothese angesehen, nach der zukünftige Kursänderungen von ihren aktuellen und vergangenen Daten abhängen. Auch weitere Untersuchungen konnten die enttäuschenden Resultate nur wenig lindern. In einer Studie von Mark (1995) [86] erzielten Strukturmodelle bei drei von vier analysierten Staaten zwar bessere Vorhersagen, allerdings nur für extrem große Prognosehorizonte von 36 und 48 Monaten. In einem erst Ende 2002 veröffentlichten Arbeitspapier erweiterten Cheung, Chinn und Pascual [23] die Untersuchungen auf lineare Wechselkursmodelle, die in den 1990er Jahren entwickelt wurden. Im Unterschied zu den genannten Studien berechneten die Autoren neben Prognosekriterien, die die Stärke der Fehlerabweichungen bestimmen, auch eine Prognosegüte für Richtungsänderungen (Sign-Prediction). Für bestimmte Prognosehorizonte und Zeitreihen waren einige Wechselkursmodelle zumindest in der Lage, signifikant bessere Richtungsänderungen als ein Random Walk zu erzeugen. Dennoch konnten auch sie keine eindeutige Unterstützung für lineare Wechselkursmodelle finden.

Der mögliche Grund für das Versagen linearer Modelle könnte in der Existenz von Nichtlinearitäten in den Finanzmarktzeitreihen liegen. Darauf deuten einige Untersuchungen wie die von Hsieh [66] (1989), Baillie & McMahon [4] (1991), Guillaume [55] (2000) oder Barnhart & Dwyer [6] (2000) hin, wobei letztere eine ganze Batterie von Testverfahren auf monatliche Zeitreihen angewendet haben. Ihre Ergebnisse basieren auf einem Bispektrum-

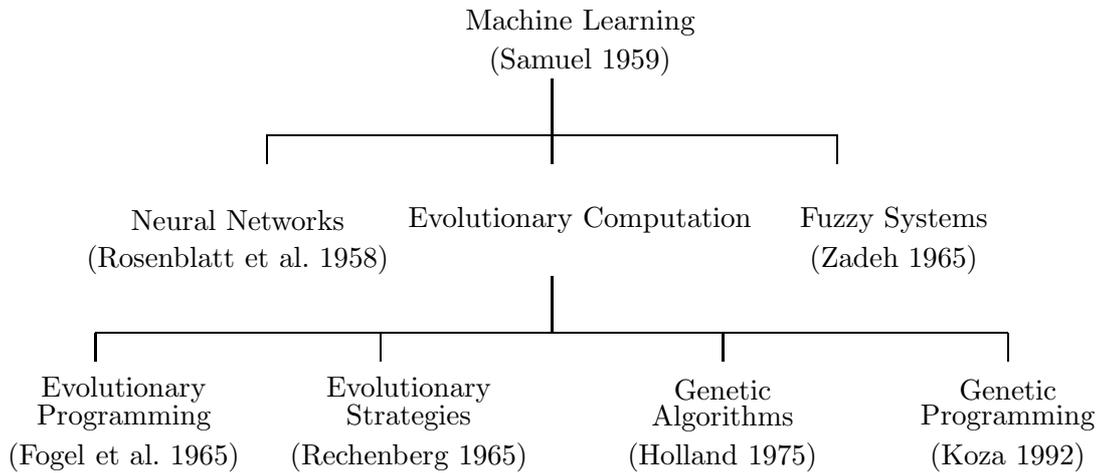


Abbildung 1.1: Forschungsbereiche, die aus dem Bereich des Maschinellen Lernens entstanden sind.

und Bikovarianztest (Hinich [62] 1982, [63] 1995), dem von Brock, Dechert und Scheinkman [14] (1987) entwickelten populären BDS Test sowie einem nichtlinearen Testverfahren von Tsay [125] (1989)¹. Abgesehen vom Bisppektraltest signalisieren alle Testverfahren eindeutig nichtlineare Eigenschaften in den Finanzmarktzeitreihen². Die für die vorliegende Arbeit verwendeten monatlichen Finanzmarktzeitreihen wurden ebenfalls auf Nichtlinearität hin analysiert, deren Ergebnisse im Anhang A dargestellt sind³. Neben den von Barnhart & Dwyer angewendeten Testverfahren wurden die Tests von McLeod & Li [88] und Engle [33] durchgeführt.

Alle Testverfahren - abgesehen vom Bisppektraltest - überprüfen die Zeitreihen auf Nichtlinearitäten, indem sie in einem ersten Schritt die linearen Abhängigkeiten durch die Anpassung eines linearen univariaten Regressionsmodells beseitigen. Die durch diesen Vorgang ('Prewhitening') erhaltenen Fehlerabweichungen u_t zwischen Beobachtungs- und geschätzten Werten dienen als Ausgangspunkt für die weitere Vorgehensweise der einzelnen Testverfahren. Da die Testverfahren die Fehlerterme auf unterschiedliche Art und Weise überprüfen, können sie bestimmte Nichtlinearitäten unterschiedlich gut nachweisen. So untersuchen die Tests von McLeod & Li und Engle die Varianz der Fehlerterme und überprüfen die Zeitreihe u_t auf ARCH Effekte. Dagegen setzen die beiden Tests von Hinich sowie die Methode von Tsay bei den Kumulanten dritter Ordnung einer stationären Zeitreihen an⁴, während das BDS Verfahren testet, ob die Schätzresiduen in einer n-dimensionalen räumlichen Anordnung Verdichtungen und somit nichtlineare Strukturen aufweisen. Folglich überprüft der BDS Test die Nullhypothese auf eine unabhängige und identische Verteilung. Der Test ist in der Lage, Nichtlinearitäten sowohl im Mittelwert als auch in der Varianz aufzudecken.

¹Nähere Erläuterungen zu den einzelnen Testverfahren befinden sich im Anhang A.

²Vgl. Barnhart [6], Tabelle 6.

³Die Schätzungen der Tabellen A.2 und A.3 erfolgten wie bei Barnhart & Dwyer mit dem Computerprogramm „Nonlinear Toolkit 4.6“ von Patterson & Ashley (2000).

⁴Die normalisierten Kumulanten dritter Ordnung entsprechen dem dritten Moment (Schiefe) einer Verteilung: Schiefe = $\frac{1}{n} \sum_t^n \frac{(x_t - \mu)^3}{\sigma^3}$ mit n als Anzahl der Beobachtungswerte der Zeitreihe x_t . Siehe auch Bouchaud [12], S.7.

Brock et al. haben in einer Monte-Carlo Studie die Güte des BDS Tests mit den Testverfahren von Engle, Tsay und dem Hinich Bispektrum anhand verschiedener dynamischer nichtlinearer Prozesse untersucht. Die Studie zeigt, daß der Engle und BDS Test die Existenz nichtlinearer Dynamiken gut aufzudecken vermögen. Dagegen versagen die Verfahren von Tsay und Hinich, die in den Varianzen der Fehlertermen enthaltenen Nichtlinearitäten bei den ARCH und GARCH Prozessen offenzulegen⁵. Beim Anwenden des BDS Tests für kleine Zeitreihen schlagen Brock et al. vor, die asymptotische Verteilung der kritischen Werte für die Nullhypothese anhand einer Bootstrap-Technik zu berechnen. Diesem Vorschlag kommt Kanzler [74] mit einem speziellen Computerprogramm nach, das zusätzlich auf die Zeitreihen angewandt wurde⁶. Die Testergebnisse stimmen mit denen von Barnhart & Dwyer überein. Demnach weisen fast alle Zeitreihen eindeutig nichtlineare Eigenschaften auf. Bei einigen Zeitreihen fallen die Signifikanzniveaus der verschiedenen Tests allerdings sehr unterschiedlich aus. Betrachtet man beispielsweise die niederländischen Inflationsraten als auffälligste Zeitreihe, so signalisieren abgesehen vom McLeod & Li Test alle Testverfahren *keine* Nichtlinearitäten in den Daten inklusive des BDS Testverfahrens, während der BDS Test nach Kanzler wiederum diese auf einem Signifikanzniveau von 0.001 deutlich untermauert. Darüber hinaus zeigt sich, daß wie bei Barnhart & Dwyer der Bispektraltest ebenfalls kaum Evidenz für Nichtlinearitäten offenlegt.

Insgesamt liefern die Testverfahren starke Indizien für das Vorliegen von Nichtlinearitäten in den Finanzmarktzeitreihen. Das Ziel der vorliegenden Arbeit ist es daher herauszufinden, wie gut genetisch entwickelte Prognosemodelle monetäre Größen vorhersagen können. Das Genetische Programmieren gehört zum Bereich des Maschinellen Lernens, aus dem eine Vielzahl an unterschiedlichen nichtlinearen Schätzverfahren hervorgegangen ist. Es ist daher sinnvoll, das Genetische Programmieren zunächst in den Kontext des maschinellen Lernens zu stellen, um die Vorteile des Genetischen Programmierens im Vergleich zu anderen populären Methoden zu verdeutlichen. Aufgrund der Komplexität einiger Forschungsbereiche wird nur eine grobe Darstellung vorgenommen.

1.1 Maschinelles Lernen

Die Wissenschaften 'maschinelles Lernen' und 'künstliche Intelligenz' bilden zusammen die tragenden Säulen der Computerwissenschaften⁷. Aus diesen beiden Bereichen haben sich mit der Zeit zahlreiche neue Forschungsbereiche entwickelt. Speziell aus dem Bereich des maschinellen Lernens sind Forschungsbereiche entstanden, die in den vergangenen Jahren durch eine Vielzahl an Veröffentlichungen im Finanzmarktbereich auf sich aufmerksam gemacht haben. Der Begriff des maschinellen Lernens wurde von Samuel [112] (1959) geprägt und verfolgt die Idee, Computerprogramme zu entwickeln, die es ermöglichen, Problem-

⁵Die Testverfahren überprüfen, ob die Kumulanten dritter Ordnung von null verschieden sind. Die ARCH und GARCH Prozesse besitzen aber eine Kumulante dritter Ordnung von null, so daß das Versagen der Tests leicht nachvollziehbar ist.

⁶Das von Kanzler geschriebene Computerprogramm für MATLAB wurde von Timur Yusupov für C++ umgeschrieben. Die in Tabelle A.1 dargestellten Ergebnisse basieren auf diesem C++ Programm.

⁷Vgl. Fogel [42].

stellungen eigenständig zu lösen, ohne den Lösungsweg vorher explizit anzugeben. Alle in Abbildung 1.1 aufgeführten Forschungsbereiche versuchen, dieses Ziel zu erreichen. Das Genetische Programmieren ist der jüngste Forschungsbereich und gehört zur Gruppe der Evolutionary Computation. Dabei diente die Natur für die Entwicklung der Evolutionary Computation wie auch die der künstlichen neuronalen Netze und der Fuzzy Systeme als Vorbild, so daß viele Begriffe aus der biologischen Terminologie stammen. Die Darstellung der einzelnen Forschungszweige erfolgt unter dem Aspekt, dem Leser die groben Funktionsweisen zu vermitteln, um ihn dadurch in die Lage zu versetzen, die wesentlichen Unterscheidungsmerkmale und somit die Vor- und Nachteile des jeweiligen Verfahrens zu erfassen⁸.

1.2 Künstliche Neuronale Netze

Eine der ersten nichtlinearen Regressionsmodelle aus dem Bereich des maschinellen Lernens sind die künstlichen neuronalen Netzwerke⁹, deren erste theoretische Ansätze bereits in den 1940er Jahren entstanden sind. Zwischen 1957 und 1958 konstruierten Frank Rosenblatt et al. am MIT den ersten erfolgreichen neuronalen Computer für Mustererkennungsprobleme. Neuronale Netze imitieren die Funktionsweise des natürlichen Nervensystems, das aus unzähligen miteinander verbundenen Nervenzellen - Neuronen - besteht. Elektrische Impulse durchströmen das Gehirn und aktivieren die Neuronen bei Überschreiten eines bestimmten Grenzwertes. Bei der Bewältigung einer Aufgabe werden schließlich ganze Bereiche bzw. Netze von Neuronen aktiviert, die untereinander interagieren. Befaßt man sich mit der Literatur über neuronale Netzwerke, so schrecken Bezeichnungen wie 'input-', 'hidden-' und 'output-layers' oder 'hidden units', 'squashing-' und 'activation function' den Leser zunächst ab und suggerieren neben etlichen weiteren Ausdrücken eine sehr komplexe Struktur. Dabei ist ein neuronales Netz aus struktureller Sicht lediglich eine nichtlineare funktionale Form, bestehend aus einer Linearkombination einzelner nichtlinearer Funktionen, wobei die Gewichte der Summanden durch das neuronale Netz optimiert werden. Die Funktionsweise kann relativ kurz dargestellt werden. Bezeichnet man mit x die Input-Variablen und mit y den Output der komplexen Funktion, dann läßt sich ein neuronales Netz folgendermaßen beschreiben:

$$y_t = \phi(\beta_0 + \sum_{i=1}^q \beta_i h_{it}) \quad (1.1)$$

mit

$$h_{it} = \psi(\gamma_{i0} + \sum_{j=1}^n \gamma_{ij} x_{jt}), i = 1, \dots, q$$

h_{it} kennzeichnen die Neuronen, q und n stehen für die Anzahl der verwendeten Neuronen und Beobachtungswerte, ψ und ϕ beschreiben die Aktivierungsfunktionen. Die

⁸Einen guten Überblick über die Vor- und Nachteile der einzelnen Forschungsbereiche liefert auch Mitze [94], so daß einige Aspekte sich an diesem Artikel orientieren.

⁹Vgl. Fogel [42], Tettamanzi et al. [122], Plasmans et al. [104].

Funktion ψ wird als “squashing function” bezeichnet, da sie die eintreffenden Werte der Input-Variablen x in Outputwerte bestimmter Bandbreite komprimiert. Die ersten ψ -Transferfunktionen erzeugen entweder den Outputwert 0 (nicht aktiviert) oder 1 (aktiviert) abhängig davon, ob die gewichtete Summe der Input Variablen x_{jt} einen bestimmten Schwellenwert überschreitet. In der neueren Forschung werden differenzierbare, in der Regel logistische ψ -Funktionen eingebaut, die Werte im Bereich zwischen Null und Eins annehmen. Bei der für die Regression relevanten ϕ -Funktion handelt es sich in der Regel ebenfalls um eine logistische Funktion. Die Parameterwerte γ und β bilden schließlich die anhand des neuronalen Netzes zu optimierenden Gewichte.

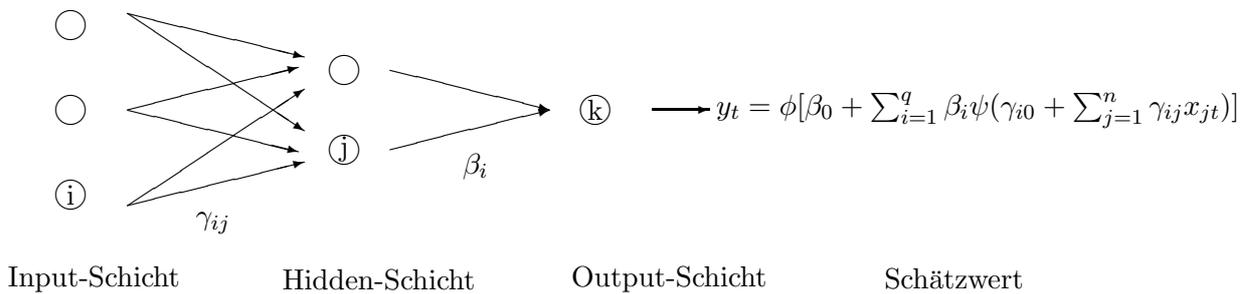


Abbildung 1.2: Vorwärtsgerichtetes Neuronales Netzwerk. Als Hidden-Layer bezeichnet man Schichten, die nicht direkt mit den Input- und Output-Werten verbunden sind.

Die Anordnung der Neuronen in einer bestimmten Struktur wird als „Netzwerk(architektur)“ bezeichnet. Abbildung 1.2 stellt das häufig verwendete sogenannte ‘feed forward neural network’ dar. Man erkennt, daß die Neuronen in verschiedene Schichten geordnet sind: der Input-, Hidden- und Output-Schicht. Soll eine Finanzmarktzeitreihe approximiert werden, fließen in die Input-Schicht die Vektoren mit den erklärenden Daten ein. Bei gegebenen Daten, der ϕ - und ψ -Funktionen errechnet das Neuronale Netzwerk durch graduelle Optimierung der Gewichte β und γ die optimale Regressionsgleichung. Als Minimierungskriterium wird üblicherweise die mittlere quadratische Fehlerabweichung (MSE) zwischen geschätzten und tatsächlichen Werten berechnet. Da der Lösungsraum mit seinen potentiellen Regressionsgleichungen allerdings häufig multi-modaler Natur mit vielen lokal minimalen Fehlerabweichungen ist, besteht aufgrund des graduellen Optimierungsverfahrens die Gefahr, das globale Minimum zu verfehlen und in einem dieser lokalen Minima stecken zu bleiben. Durch wiederholte Anwendungen wird versucht, dieser Problematik entgegenzuwirken. Ein anderes Problem stellt das Overfitting dar, das auf der Suche nach der besten Schätzgleichung durch folgende Erscheinung gekennzeichnet ist: im Zeitablauf verringert sich die Fehlerabweichung kontinuierlich, und die Regressionsgleichung bildet die zu schätzenden Daten immer besser ab. Gleichzeitig verschlechtern sich ab einem bestimmten Zeitpunkt jedoch die Prognosegüte der Regressionsgleichung.

Das neuronale Netzwerk beginnt, die zufälligen Schwankungen zu stark zu berücksichtigen bei gleichzeitiger Vernachlässigung des zugrunde liegenden Trends oder der strukturellen Zusammenhänge zwischen den Variablen. Während das Overfitting Problem ein allgemeines Problem nichtlinearer Modelle darstellt, ist das als 'Black-Box' bezeichnete Verhalten ein spezifisches Problem neuronaler Netze und beschreibt den Sachverhalt, das neuronale Netze bei Vorliegen guter Ergebnisse kaum analytische Interpretationen zwischen Eingabe- und Ausgabewerten zulassen¹⁰. Ein weiterer häufig in der Literatur erwähnter Kritikpunkt besteht darin, daß die Optimierung der Gewichte lediglich eine Erweiterung aber keine wesentliche Neuerung konventioneller Regressionsmethoden darstellt¹¹. Weiterhin wird beanstandet, daß künstliche neuronale Netze einen langsamen Lernprozeß in Abhängigkeit von der Größe der Netzwerkarchitektur und der Größe der zu schätzenden Datenmenge aufweisen¹².

1.3 Fuzzy System

Der Forschungsbereich der Fuzzy Systeme spielt bei Anwendungen auf den Finanzmärkten eher eine untergeordnete Rolle¹³. Allerdings existieren in der Literatur einige Studien, die die Fuzzy Systeme mit anderen Bereichen des maschinellen Lernens kombiniert haben. Entwickelt wurde dieses Verfahren von Lotfi A. Zadeh (1965) an der University of California. Obwohl es in den Vereinigten Staaten erfunden worden ist, fanden Fuzzy Systeme zunächst vor allem in Japan große Beachtung. Das Verfahren versucht, die Art und Weise des menschlichen Denkens nachzuahmen. Charakteristisch für das menschliche Denken ist es, komplexe Umweltzustände anhand der verwendeten Linguistik nicht eindeutig beschreiben zu können. Im Gegensatz zu künstlichen Neuronalen Netzen, deren ursprünglichen Aktivierungsfunktionen nur die beiden Extremzustände 'aktiviert' und 'nicht-aktiviert' zulassen, erlaubt das Fuzzy System einen fließenden, verschwommenen (=fuzzy) Übergang bei der Bewertung von Zuständen. Überträgt man die Idee der Fuzzy Systeme auf Finanzmärkte bedeutet dies, daß ein neuronales Netzwerk eine Inflationsrate beispielsweise bis 2% Prozent als niedrig, darüber als hoch und ab 3% als sehr hoch definiert. Bei Finanzmarktuntersuchungen kann es aber wünschenswert sein, einen fließenden Übergang zu betrachten, um eine Inflationsrate von 2.0% als niedrig und eine Inflationsrate von 2.01% nicht bereits als hoch einzustufen. Diese Aufgabe übernehmen die Fuzzy Systeme. Die Funktionsweise eines Fuzzy Systems läßt sich am einfachsten durch ein kurzes Beispiel erläutern. Die Zentralbank soll die Zinssätze in Abhängigkeit des Konjunkturwachstums und der Inflationsrate regulieren. Dazu definiert sie für die Inflationsraten eine Fuzzy Menge $\{F_i^{Infl} | i = 1, 2, 3\}$, der die linguistischen Attribute, *niedrig*, *hoch*, *sehr hoch* zugeordnet werden. Der Grad der Zugehörigkeit der Inflationsrate wird durch eine Funktion bestimmt, deren Werte zwischen 0 und 1 liegen. Zum Beispiel könnte eine Inflationsrate von 2.5% mit dem Grad 0.7 zur Fuzzy Menge *hoch* und mit dem Grad 0.15 zur Fuzzy Menge *sehr hoch* gehören. Für die anderen Zeitreihen werden ähnliche Fuzzy Mengen erstellt. Nach

¹⁰Vgl. Swanson & White [119], S. 542 oder Lin [78], S. 7.

¹¹Vgl. Cheng & Titterington [19] oder Tettamanzi [122], S. 81.

¹²Vgl. Moshiri & Cameron [17], S. 11.

¹³Vgl. Fogel, D. [42] (1995).

der Erstellung der Fuzzy Mengen formuliert der Anwender sein Wissen in Form einfacher *WENN-DANN*-Regeln: *WENN Inflation = hoch UND Wachstum = sehr hoch DANN Zinsen = hoch*. In Abhängigkeit der Kombinationsmöglichkeiten muß der Anwender eine entsprechende Anzahl an Regeln aufstellen. Liegen der Zentralbank nun konkrete Zahlen für die Inflations- und Wachstumsrate vor, dann kann anhand der Zugehörigkeitsgrade der Erfülltheitsgrad einer Regel und durch weitere Berechnungen ein konkreter Vorschlag zur Änderung der Zinsen ermittelt werden.

Ein Nachteil der zumindest ursprünglichen Fuzzy Systeme ist der fehlende Lernalgorithmus, der Korrekturen bei schlechten Handlungsempfehlungen vornimmt. Goonatilake bemängelt, daß die Funktionen für die Zugehörigkeitsgrade und das Expertenwissen im vornherein festgelegt werden müssen. Neben dem damit verbundenen Zeitaufwand stellen sie mögliche Quellen für eine spätere Modellfehlspezifikation dar¹⁴. Der Vorteil durch die a priori Implementierung von Expertenwissen mit Hilfe der Wenn-Dann-Regeln liegt in einem transparenten Entscheidungsprozeß, der im Gegensatz zur Black-Box Eigenschaft der künstlichen neuronalen Netze nachvollziehbar ist und Interpretationen zuläßt. Aufgrund ihrer Vorteile werden Fuzzy Systeme in Kombination mit anderen Forschungsbereichen angewendet.

1.4 Evolutionary Computation

Ein weiterer großer Forschungszweig, der aus dem Bereich 'Maschinelles Lernen' entstanden ist, sind die evolutionären Algorithmen. Der Begriff steht für eine Gruppe von Suchverfahren, die auf der Konzeption des natürlichen Evolutionsprozesses von Charles Darwin (1859) [25] beruhen. Hierzu gehören die Forschungsbereiche Evolutionary Programming (Fogel et al., 1966), Evolutionary Strategies (Rechenberg [107], 1965 und Schwefel [114], 1965), Genetische Algorithmen (Holland [64], 1975) und das Genetische Programmieren (Koza [75], 1992). Die Theorie von Darwin beruht auf dem Prinzip des "survival of the fittest" und besagt, daß nur die stärksten Organismen im Zeitablauf überleben und sich weiterentwickeln. Dieser Prozeß der Evolution durch Selektion ist allen Forschungsbereichen gemein und spiegelt sich in einem einheitlichen Programmablauf wider. Demnach generieren alle Techniken eine zufallsbedingte Anfangsgeneration, die eine bestimmte Anzahl möglicher Lösungen enthält. Die Güte der Lösungskandidaten wird mit Hilfe eines Fitnesskriteriums wie beispielsweise der mittleren quadratischen Fehlerabweichung bewertet und entsprechend geordnet. Evolutionsoperatoren produzieren anhand der besten Kandidaten neue Lösungen, die erneut bewertet, geordnet und selektiert werden. Das Programm wiederholt die Generierung und Selektion neuer Populationen solange, bis ein vorher festgelegtes Abbruchkriterium erfüllt wird.

1.4.1 Evolutionsstrategien

Die Evolutionsstrategien (ES) wurden von den beiden Autoren Schwefel und Rechenberg unabhängig voneinander an der Technischen Universität Berlin entwickelt. Ursprünglich

¹⁴Vgl. Goonatilake [51] S. 17.

lag das Anwendungsgebiet der Evolutionsstrategien auf dem Gebiet der Parameteroptimierung. Ein Beispiel für die Berechnung optimaler Parameter liefert der Aufsatz von Binner et al. (2004) [9], die anhand unterschiedlicher Geldmengenaggregate Inflationsprognosemodelle aufstellten und mit Hilfe der Evolutionstrategien die optimalen Regressionkoeffizienten berechneten. Die Individuen eines ES-Programms werden durch zwei reelle Zahlenvektoren $v = (x, \delta)$ repräsentiert, wobei der erste Vektor x einen Lösungsvorschlag und der zweite Vektor Standardabweichungen enthält¹⁵. Obwohl mittlerweile auch Crossover bzw. Rekombinationen als genetische Operationsmethode berücksichtigt werden, bilden die Mutationsoperationen den Schwerpunkt bei der Schaffung neuer Lösungskandidaten. Demnach wird bei einer Mutation ein Wert aus dem Vektor x durch Addition einer normalverteilten Zufallszahl verändert:

$$x_{t+1} = x_t + N(0, \delta) \tag{1.2}$$

Das Ausmaß der Veränderung hängt von der Standardabweichung δ des zweiten Vektors ab, die wiederum in Abhängigkeit von der Fehlerabweichung zur Zielvariablen steht. Größere Fehlerabweichungen verursachen daher aufgrund einer größeren Standardabweichung größere Mutationsschritte mit einer größeren Wahrscheinlichkeit und umgekehrt. Durch die Endogenisierung erhält man einen Lernprozeß, bei dem sich die Standardabweichung dynamisch anpaßt. Zusätzlich treten aufgrund normalverteilter Zufallszahlen unabhängig vom Anpassungsprozeß der Standardabweichung kleinere Veränderungen ebenfalls häufiger auf als große, so daß insgesamt betrachtet der Vorgang dem natürlichen Evolutionsprozeß entspricht. Die Anfangspopulation enthält eine bestimmte Anzahl an μ zufällig erzeugten Individuen $v(x, \delta)$. Für die Bildung neuer Populationen existieren zwei Verfahren, die mit $(\mu + \lambda)$ und (μ, λ) bezeichnet werden. Beide Methoden generieren eine Anzahl an λ neuen Individuen. Während das erste Verfahren die neuen Lösungskandidaten aus der Menge der alten *und* neuen Individuen $(\mu + \lambda)$ auswählt, erfolgt dies beim zweiten Verfahren lediglich aus der Menge λ , wobei die Auswahlmenge größer als die eigentliche Population ist ($\lambda > \mu$). Die Selektion der Lösungskandidaten für genetische Operationen ist abhängig von deren Güte. Als Güte- bzw. Fitnesskriterium kann beispielsweise die bereits bei den neuronalen Netzen erwähnte mittlere quadratische Fehlerabweichung verwendet werden. Je geringer die Fehlerabweichung, desto größer ist die Wahrscheinlichkeit für genetische Operationen ausgewählt zu werden. Die Generierung einer neuen Population durch genetische Operationen, Bewertung und Ordnung der einzelnen Lösungskandidaten und Auswahl einzelner Lösungskandidaten für erneute genetische Operationen erfolgt solange, bis ein vorher festgelegtes Abbruchkriterium erfüllt wird.

1.4.2 Evolutionary Programming

Während das Motiv der ES im Bereich der Parameteroptimierung anzusiedeln ist, wird der Aufgabenbereich des Evolutionary Programming (EP) in der Literatur häufig mit "Intelligence Behaviour" umschrieben, wobei sich hinter diesem Begriff häufig die Aufgabe der

¹⁵Vgl. auch Pedrycz [102].

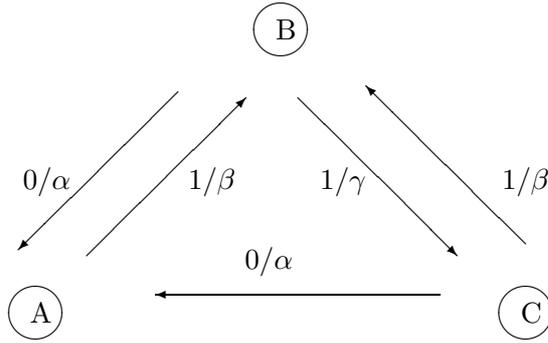


Abbildung 1.3: Eine Finite State Machine mit drei Zuständen. Zustandsüberprüfungen sind links und die Pointer rechts vom Schrägstrich abgebildet.

Mustererkennung verbirgt. Dieser Suchalgorithmus wurde im Jahr 1965 von Fogel, Owens und Walsh [43] dargestellt. Betrachtet man beispielsweise eine Finanzmarktzeitreihe über einen bestimmten Zeitraum, könnte sich ein Investor die Frage stellen, ob der Kursverlauf häufig wiederkehrende Muster aufweist, anhand derer man mit einer bestimmten Wahrscheinlichkeit in der Lage wäre, Vorhersagen zu treffen. Ein zu untersuchendes Muster könnte folgende Kursentwicklung im Zeitablauf t sein:

1. Zeitintervall $t, t + 1$: Kurs P steigt (Zustand A)
2. Zeitintervall $t + 1, t + 2$: Kurs P stagniert (Zustand B)
3. Zeitintervall $t + 2, t + 3$: Kurs P fällt (Zustand C)
4. Zeitintervall $t + 3, t + 4$: Kurs P stagniert (Zustand B)

Um diese Kombination ABCB der drei verschiedenen Zustände im Hinblick auf die Zeitreihe zu untersuchen, bedient sich das EP sogenannter Finite State Machines (FSM) [102]. Sie sind das Hauptmerkmal der EP und stellen aus technischer Sicht kleine Computerprogramme dar. Schaubild 1.3 verdeutlicht den Ablauf einer FSM. Den Ausgangszustand bildet bei dieser Kombination der Zustand A bei einem gegebenen Kurswert P_t (Ausgangswert) zum Zeitpunkt t . Bekommt das Computerprogramm als neue Inputinformation den Kurswert im Zeitpunkt $t + 1$, überprüft es die Richtungsänderung und vergleicht es mit dem vorgegebenen Schema ABCB. Ist der Kurs gestiegen ($P_{t+1} > P_t$), so stimmt die Richtungsänderung mit Zustand A überein ($1 = \text{wahr}$) und die Maschine wechselt in den Zustand B¹⁶. Anderenfalls ($0 = \text{falsch}$) verharrt es im Zustand A und der Kurs zum Zeitpunkt $t + 1$ dient als neuer Ausgangswert. Unter der Annahme, daß der Kurs gestiegen und die FSM infolgedessen in den Zustand B gesprungen ist, bekommt die FSM als neue Inputinformation den Kurs P_{t+2} . Hat sich der Kurs nicht geändert ($P_{t+2} = P_{t+1}$), wechselt das Programm in den Zustand C, ansonsten springt es in den Zustand A zurück

¹⁶Die Pointer α, β, γ bestimmen, in welchen Zustand das Computerprogramm wechseln soll.

und bekommt als neuen Ausgangswert P_{t+3} usw. Jede komplette Übereinstimmung des Schemas mit der tatsächlichen Kursentwicklung wird gespeichert und am Schluß zur Bewertung der FSM herangezogen. Zu Beginn des Programms wird eine bestimmte Anzahl solcher FSM's mit zufälligen Schemata generiert, die die Anfangspopulation bilden. Die genetischen Operationen beschränken sich auf Mutationsoperationen und setzen direkt an den FSM's an, in dem sie zum Beispiel Zustände hinzufügen oder entfernen. Das Ausmaß der Mutationsänderungen kann abhängig von der Problemstellung wie bei der ES dynamisch erfolgen und paßt sich abhängig von der Fehlerabweichung an. Für die Selektion wird im Allgemeinen ein stochastisches Tournamentverfahren benutzt. Bei diesem Verfahren steht jeder ausgewählte Lösungskandidat (FSM) im Wettbewerb mit einer bestimmten Anzahl anderer Lösungskandidaten, so daß durch den relativen Vergleich auch schlechtere FSM's selektiert werden können. Somit unterscheidet sich dieses Auswahlverfahren von der deterministischen Methode der EP.

1.4.3 Genetische Algorithmen

Die Genetischen Algorithmen (John Holland, 1975) bilden zusammen mit den neuronalen Netzen die beiden populärsten Forschungsdisziplinen im Bereich computerbasierter Suchalgorithmen für Finanzmarktanwendungen. Das Hauptmerkmal der Genetischen Algorithmen liegt in der Kodierung des Lösungsraumes, die ursprünglich in Form binärer Zahlenstränge mit einer fixen Länge erfolgte. Mit einem Zahlenstrang bestehend aus zwei binären Ziffern lassen sich beispielsweise vier Kombinationen $(0, 0)$, $(1, 1)$, $(1, 0)$, $(0, 1)$ bilden, anhand derer verschiedene Handelsstrategien verschlüsselt oder im Falle einer Parameteroptimierung vier unterschiedliche Werte eines Koeffizienten überprüft werden könnten. Dabei wächst die Anzahl der Kombinationsmöglichkeiten mit dem Faktor 2^i , wobei i die Anzahl der Ziffern eines Zahlenstranges kennzeichnet. Die beiden Beispiele lassen die Probleme Genetischer Algorithmen bereits erahnen. Bei der Untersuchung von Handelsstrategien oder ähnlichen Problemstellungen muß jeder Kombinationsmöglichkeit eine konkrete Eigenschaft vom Anwender zugeordnet werden, was aufgrund des exponentiellen Wachstumsfaktors einen erheblichen Zeitaufwand verursacht. Darüber hinaus wird die Güte Genetischer Algorithmen häufig von der ex-ante Auswahl der Eigenschaften beeinflusst¹⁷. Das Problem der Parameteroptimierung liegt aufgrund der diskreten Änderungsschritte in den nur annähernd optimalen Ergebnissen. Ein Vorteil der Kodierung besteht dagegen in der leichten Anwendbarkeit genetischer Operationen, wie aus der Abbildung 1.4 hervorgeht, wobei Crossoveroperationen gegenüber Mutationen wesentlich stärker gewichtet werden. Während bei der Mutation lediglich eine Ziffer konvertiert wird, tauscht der Crossoveroperator aus zwei verschiedenen Zahlensträngen zwei beliebig abgeschnittene Teilstränge gleicher Längen miteinander aus. Wie in den evolutionären Algorithmen zuvor erstellt das Programm zunächst anhand eines Zufallsgenerators eine Anfangspopulation bestehend aus einer bestimmten Anzahl an Zahlensträngen. Die anschließende Generie-

¹⁷Vgl. Walker et al. [128], S. 52. Goonatilake [51] betont, daß die Auswahl durch vorgeschaltete Simulationsläufe optimiert werden müssen (S.14).

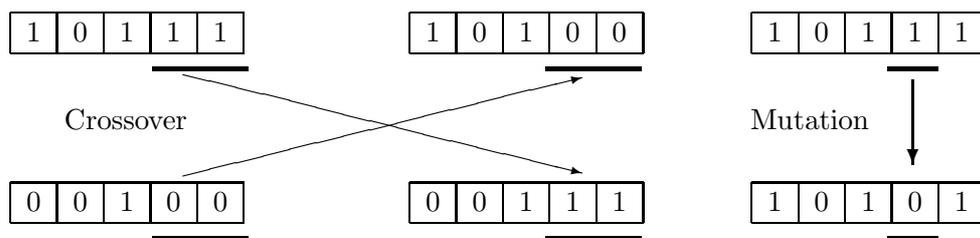


Abbildung 1.4: Evolutionsoperatoren bei Genetischen Algorithmen

rung und Selektion neuer Lösungskandidaten in den Folgeperioden erfolgt solange, bis ein zuvor festgelegtes Abbruchkriterium erfüllt ist. Für die Selektion steht die Menge der alten und neuen Individuen zur Verfügung, wobei in der Literatur verschiedene Selektionsverfahren existieren. Im Vergleich zu den künstlichen neuronalen Netzwerken konvergieren Genetische Algorithmen schneller zu einem Optimum. Zudem ist die Wahrscheinlichkeit, in einem lokalen Minimum stecken zu bleiben, gegenüber dem graduellen Optimierungsverfahren der künstlichen neuronalen Netzwerke geringer¹⁸.

1.5 Anwendungen computerbasierter Suchalgorithmen auf den Finanzmärkten

Der rasante technologische Fortschritt bei der Entwicklung von Computern und deren Rechengeschwindigkeit erleichterte die Anwendung computerbasierter Suchalgorithmen mit der Folge, daß die Anzahl der Publikationen insbesondere im Bereich der Finanzmärkte seit einigen Jahren einen regelrechten Boom erfahren hat. Die finanzmarkttheoretischen Anwendungen umfassen die verschiedensten Themengebiete, angefangen von Prognosemodellen, dem Entwickeln von Handelsstrategien, Portfoliooptimierung oder Optionspreisbestimmungen bis hin zu spieltheoretischen Entscheidungsmodellen. Mittlerweile existieren für die jeweiligen Forschungsdisziplinen mehrere Sammelwerke mit Zusammenstellungen finanz- und kapitalmarkttheoretischer Anwendungen. Im Bereich künstlicher neuronaler Netze sind die Publikationen von Refenes et al. [108], Trippi [124] sowie eine bibliographische Übersicht von Adya und Collopy [1] erwähnenswert, während Ribeiro et al. [109], Gil-Aluja [49] und Pedrycz [102] (S. 299 ff.) einen guten Überblick für Fuzzy Systeme geben. Im Hinblick auf Evolutionary Computation sind vor allem die beiden Bücher von Chen [21] [20] zu nennen. Abschließend sollen die Publikationen von Chen [22] und Gooatilake et al. [51] angeführt werden, die einen Querschnitt empirischer Untersuchungen zu den hier vorgestellten Forschungsbereichen geben.

Die Ergebnisse der nichtlinearen Schätzmethoden fallen sowohl zwischen als auch innerhalb der verschiedenen Anwendungsgebiete durchgewachsen aus. So sind auf dem Gebiet der Wechselkursprognosen einige Studien - vorwiegend Anwendungen künstlicher neuronaler Netzwerke - durchgeführt worden, deren Aussagen über die Vorteilhaftigkeit nichtlinearer Schätzverfahren deutlich voneinander abweichen. Dabei wurden die anfangs geweckten Hoffnungen, daß diese Modelle eine höhere Prognosequalität aufweisen, durch die Studie von Diebold & Nason [29] (1990) zunächst gedämpft. Die Autoren berechneten für

¹⁸Vgl. Takagi [120], S. 8-9.

wöchentliche Wechselkursraten zehn verschiedener Währungen zwischen Januar 1973 und September 1987 für ihre univariaten und rekursiv ermittelten nichtlinearen Regressionsmodelle verschiedene k-step-ahead Out-of-Sample Punktprognosen mit dem Ergebnis, daß ihre Modelle weitestgehend den Prognosequalitäten eines naiven Random Walk Modells unterlegen waren.

In Kontrast zu den Ergebnissen von Diebold & Nason steht die Studie von Kuan & Liu (1995) [77], die für tägliche Wechselkursrenditen des Zeitraumes März 1980 bis Januar 1985 Richtungs- und Ein-Monats-Punktprognosen durchführten und zumindest zu einem gewissen Ausmaß Unterstützung für die Vorteilhaftigkeit nichtlinearer Schätzverfahren entdecken konnten. Dabei beruhen die Ergebnisse von Kuan & Liu auf der nichtlinearen Prognosemethode künstlicher neuronaler Netze, während Diebold & Nason das sogenannte Neigherst Neighbor Verfahren¹⁹ angewendet haben. Insgesamt zeigen ihre Modelle bei zwei von fünf untersuchten Wechselkurszeitreihen signifikant bessere Richtungs- und Punktprognosen.

Eine weitere empirische Untersuchung, die die ermutigenden Resultate von Kuan & Lin unterstützen, wurde von Episcopos [35] (1996) durchgeführt. Seine Resultate basieren ebenfalls auf dem Suchalgorithmus künstlicher neuronaler Netze und dem Verwenden täglicher Wechselkursrenditen. Im Gegensatz zu den beiden Studien zuvor, die den U.S.-amerikanischen Dollar als Ausgangswährung für verschiedene bilaterale Wechselkursraten zugrunde legten, verwendet Episcopos den kanadischen Dollar und betrachtet den Zeitraum von Januar 1992 bis Dezember 1994. Der Autor berechnet für verschiedene künstliche neuronale Netze Ein-Monats-Punktprognosen und vergleicht die Ergebnisse mit denen eines Random Walks und eines EGARCH Modells. Seine Untersuchungen zeigen, daß sowohl das EGARCH Modell als auch die künstlichen neuronalen Netze für alle fünf Wechselkurszeitreihen den Random Walk schlagen, und daß die künstlichen neuronalen Netze im direkten Vergleich zum EGARCH Modell teilweise besser abschneiden.

Ein Jahr nach der Veröffentlichung der Resultate von Episcopos erschien eine Studie von Rauscher [106], der für die monatliche DM/U.S.-\$ Wechselkurszeitreihe des Zeitraumes Januar 1986 bis März 1997 Ein-Monats-Punktprognosen untersucht und seine anhand künstlicher neuronaler Netze geschätzten Ergebnisse mit einem Vektorfehlerkorrekturmodell (VECM) und einem Random Walk vergleicht. Im Unterschied zu den bisher vorgestellten univariaten Regressionen schätzt er multivariate Modelle auf der Grundlage eines flexiblen monetären Preismodells nach Frenkel und Bilson. Als Inputvariablen des neuronalen Netzes dienen die Variablen des ermittelten VEC-Modells. Rauscher erstellt insgesamt fünf verschiedene Prognosekriterien, unter anderem eine um die Anzahl der verwendeten Regressoren korrigierte Punktprognose (Theils Kriterium) und Richtungsprognosen. Die nichtlinearen Modelle erzielen sowohl bessere Prognosen als ein VECM als auch - in den meisten Fällen - im Vergleich zu einem Random Walk. Allerdings werden keine Signifikanztests durchgeführt.

Die Artikel von Plasmans et al. [104] aus dem Jahr 1998 bzw. Verkooijen et al. [127] (1996) gehen in die gleiche Richtung wie die Untersuchungen von Rauscher. Ihre Untersuchungen

¹⁹Zum Neigherst Neighbor Verfahren, das eine lokal-gewichtete Regressionmethode darstellt, siehe auch Banzhaf et al. [5], S. 19 ff. und Diebold & Nason [29].

sind allerdings umfangreicher und lehnen sich stärker an den Artikel von Meese & Rogoff an. Die Autoren untersuchen neben dem flexiblen monetären Preismodell von Frenkel [45] und Bilson [8] das monetäre Modell mit kurzfristig starren Preisen von Dornbusch [30] und Frankel [44] sowie ein portfoliotheoretisches Modell für monatliche Wechselkurszeitreihen von Januar 1974 bis Juli 1994. Für die betrachteten Niveauvariablen konnten sie für verschiedene Prognosehorizonte weder lineare noch nichtlineare langfristige Beziehungen zwischen den fundamentalen ökonomischen Variablen entdecken, so daß die anhand künstlicher neuronaler Netze erstellten Strukturmodelle weder im Vergleich zu einem Random Walk noch zu nichtlinear geschätzten univariaten Zeitreihenmodelle bessere Vorhersagen erzeugten. Lediglich die kurzfristigen Ein-Monats-Prognosen differenzierter Zeitreihen lieferten Hinweise auf eine schwache fundamentale Struktur, allerdings sowohl für lineare als auch für nichtlineare Modelle. Auf die letzte Studie wird in einem späteren Kapitel dieser Arbeit noch einmal näher eingegangen.

Die computerbasierten Suchalgorithmen haben nicht nur zahlreiche theoretische und empirische Untersuchungen auf den unterschiedlichsten Gebieten hervorgebracht, sondern sie haben neben der erfolgreichen Etablierung auf den jeweiligen Anwendungsgebieten auch in den wirtschaftlichen Entscheidungsprozessen der Unternehmen Einzug erhalten. Das Buch von Goonatilake beschäftigt sich mit praktischen Anwendungen, die in das Tagesgeschäft von Unternehmen erfolgreich integriert wurden. Einige Anwendungen aus diesem Buch werden im Folgenden kurz vorgestellt, um dem Leser zu verdeutlichen, daß computerbasierte Suchalgorithmen nicht nur theoretische Konstrukte sind, sondern deren Nutzen konkret in Geldeinheiten ausgerechnet werden kann. Das erste Beispiel stellt ein Programm gegen den Mißbrauch von Kreditkarten vor, das auf der Basis künstlicher neuronaler Netze konzipiert worden ist. Das als *Cardholder Risk Identification Service* (CHRIS) bezeichnete Computerprogramm erstellt für jeden Kartenbesitzer ein für ihn typisches Verhaltensprofil und überprüft jede eingehende Transaktion auf Verhaltensabweichungen. Eine solche Software wurde bei dem bekannten Kreditunternehmen Visa International integriert mit dem Ergebnis, daß bereits in den ersten sechs Monaten Kosten in Höhe von ungefähr 40,- Mio. U.S.-\$ eingespart werden konnten.

Auf dem Gebiet der Kreditgeschäfte wurde von einem italienischen Kreditinstitut ein auf Genetischen Algorithmen basierendes Computerprogramm eingeführt, das mit Hilfe von Bilanzrelationen die Gefahr einer Zahlungsunfähigkeit von Unternehmen vorhersagen soll. Das Programm erhielt bei dem Kreditinstitut gegenüber einer auf der Basis künstlicher neuronaler Netzwerke entwickelten Software den Vorrang, da die Gründe für eine mögliche Zahlungsunfähigkeit eines Unternehmens bei Genetischen Algorithmen im Gegensatz zu künstlichen neuronalen Netzwerken mit ihrer Black-Box Eigenschaft einsehbar sind.

Im Bereich der Fuzzy Systeme wurde ein intelligentes Handelsprogramm für kurzfristige Anleihen entwickelt, das die japanische Fuji Bank in ihr Tagesgeschäft eingebunden hat. Das Anfang der 1990er Jahre eingeführte Programm bestand seinerzeit aus mehr als 200 Fuzzy Regeln auf der Grundlage bekannter technischer Handelsstrategien. Da am Finanzmarkt profitable Regeln schnell erkannt und absorbiert werden, müssen die Händler die Menge an Fuzzy Regeln kontinuierlich mit neuen Strategien erweitern. Die Software unterstützt die Händler, indem sie ihre neu definierten Regeln mit Hilfe historischer Da-

ten auf ihre Profitabilität hin überprüfen können. Zwei Jahre nach Inbetriebnahme des Programms beliefen sich die Einnahmen gemäß Goonatilake auf ungefähr 770000,- U.S.-\$ monatlich.

Als letztes soll ein Beitrag von Didner [26] erwähnt werden, der die erfolgreiche Einführung eines Systems basierend auf Expertenwissen beim U.S. Unternehmen American Express im November 1988 schildert. Bis zur Einführung des Computerprogramms erfolgte die Kreditvergabe an Privatkunden nach einem relativ starren System unter Zuhilfenahme eines bevollmächtigten Angestellten. Die Entscheidung über die Zulässigkeit und Höhe der Kreditvergabe traf der Angestellte anhand vorliegender vom Privatkunden eingeforderter Informationen. Bei einer zu leichtfertigen Vergabe von Krediten oder bei unterlassenen Kreditzusagen kann dem Unternehmen aufgrund von Zahlungsunfähigkeit eines Kunden bzw. durch ein entgangenes profitables Geschäft ein Nachteil entstehen.

Die Ziele des Computerprogramms bestanden darin, die Entscheidungsfindung über Zu- oder Absage und Höhe des Kredites zu beschleunigen und anhand von Expertenwissen erfahrener und besonders erfolgreicher Angestellter Entscheidungsregeln für das Computerprogramm zu erstellen, um zu riskante Kreditgeschäfte ex-ante besser einschätzen zu können. Diese Zielvorgaben konnte das Programm erfolgreich erfüllen und wurde im November 1988 nach einer dreijährigen Erprobungsphase in Betrieb genommen. Neben der schnelleren Kreditabwicklung generierte das Programm aufgrund einer differenzierteren Entscheidungsfindung bei riskanten Kreditgeschäften zusätzliche, profitable Kreditsätze. Die Kosten pro Transaktion konnten durch das Programm deutlich von 15,- U.S.-\$ auf 1,40 U.S.-\$ gesenkt werden.

1.6 Zusammenfassung

Die zu Beginn durchgeführten Untersuchungen der verwendeten Finanzmarktzeitreihen auf Nichtlinearität anhand verschiedener Testverfahren unterstützen die Vermutung, daß die Finanzmarktzeitreihen auch bei Vorliegen niedrig frequenter monatlicher Daten nichtlineare Eigenschaften besitzen. Die Ergebnisse dienen als Motivation für die Anwendung eines nichtlinearen Schätzverfahrens auf monetäre Größen. Neben dem Genetischen Programmieren existieren weitere nichtlineare Schätzverfahren, die alle dem Forschungsbereich des maschinellen Lernens zuzuordnen sind. Die Konzepte ausgewählter nichtlinearer Schätzverfahren wurden kurz vorgestellt, um die Vorteile des Genetischen Programmierens, das im folgenden Kapitel behandelt wird, besser nachvollziehen zu können.

Kapitel 2

Genetisches Programmieren

Als John Koza [75] im Jahr 1992 sein Buch “Genetic Programming” publizierte, sorgte er unter diesem Namen für die Etablierung des jungen Forschungsbereichs auf dem Gebiet der Evolutionary Computation. Allerdings charakterisieren viele Autoren in der Literatur das Genetische Programmieren vor allem wegen der Aufhebung der Kodierung häufig auch als schlichte Weiterentwicklung der Genetischen Algorithmen. Dabei bedeutet der Wegfall dieser Restriktion nicht nur eine größere Flexibilität im Hinblick auf einen theoretisch unbegrenzten Lösungsraum oder der freieren Gestaltungsmöglichkeit bei den genetischen Evolutionsoperatoren, sondern sorgt wegen der fehlenden - durch den Anwender vorzunehmenden - Verschlüsselung auch zu einer leichteren Handhabung. Aufgrund unkodierter Lösungskandidaten setzen die genetischen Operationen nicht bei den Binärsträngen bestehend aus Nullen und Einsen sondern direkt bei den mathematischen Termen einer Gleichung an.

Koza verwendet die Computersprache LISP, um mathematische Gleichungen in einer anderen Syntax wiederzugeben. Damit lassen sich die Lösungen in Form einer Baumstruktur abbilden. Der Vorteil dieser Repräsentation liegt vor allem in der leichteren Anwendung der genetischen Operationen, die in Abbildung 2.1 mit dem Ablauf eines Crossovers und in Abbildung 2.2 mit dem Ablauf einer Mutation graphisch erklärt werden. Die Funktion $Y * X + \text{Sinus}(Z)$ wird in LISP folgendermaßen ausgedrückt: $(+(*YX)(\text{Sinus}Z))$. Durch die Darstellung der mathematischen Funktion als Baumstruktur erhält man interne Knotenpunkte mit Funktionsoperatoren und Endknotenpunkte mit Konstanten und Variablen. Wie beim Genetischen Algorithmus werden auch hier zwei Elternfunktionen benötigt, zwischen denen jeweils ein zufällig ausgewählter Teilbereich ausgetauscht wird. Die Teilbereiche dürfen nicht willkürlich abgetrennt werden, sondern unterliegen bestimmten Regeln. Entweder wird an einem Endknotenpunkt abgeschnitten und die Konstante oder Variable dient als Teilbereich oder es wird an einem internen Knotenpunkt angesetzt. In diesem Fall bildet der Knotenpunkt zusammen mit dem gesamten Strang unterhalb dieses internen Knotens den auszutauschenden Teilbereich.

Das dieser Arbeit zugrunde liegende Genetische Programm wurde mit dem Matrizenrechenprogramm GAUSS für Windows NT/95 Version 3.2.38 geschrieben und soll daher im weiteren mit GP GAUSS bezeichnet werden. Die LISP Schreibweise findet hier keine Berücksichtigung.

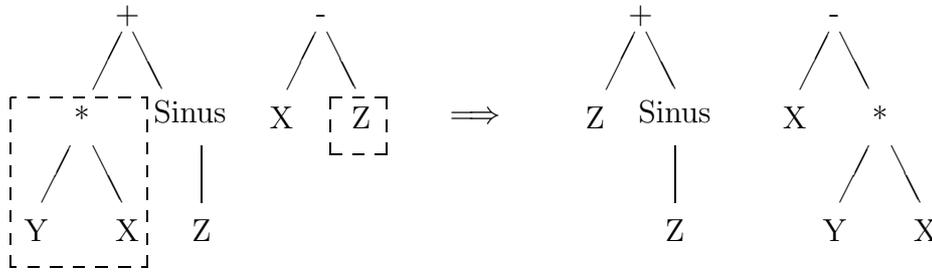


Abbildung 2.1: Crossover: Die Darstellung der mathematischen Funktionen in Form einer Baumstruktur vereinfacht die Anwendung der genetischen Operationen; sie werden von unten nach oben sowie von links nach rechts gelesen.

Um die Baumstruktur und einen korrekten Vorgang der genetischen Operationen dennoch zu gewährleisten, wurden sämtliche Operanden und Operatoren bis auf Addition, Subtraktion, Division und Multiplikation in Klammern gesetzt. Die oben erwähnte Funktion sähe in diesem Programm wie folgt aus: $(Y) * (X) + (\text{Sinus}(Z))$. Darüber hinaus setzt das Programm weitere Klammern anhand eines Zufallsprozesses¹. Sollen Teilbereiche einer Funktion abgeschnitten und ausgetauscht werden, wählt das Programm zufällig eine öffnende Klammer aus und sucht die entsprechende schließende Klammer, so daß der auszutauschende Teilbereich für sich genommen einen abgeschlossenen mathematischen Term darstellt. In diesem Beispiel hätte das Programm zwei zusätzliche Klammern setzen müssen, damit der oben abgebildete Crossovervorgang möglich gewesen wäre: $((Y) * (X)) + (\text{Sinus}(Z))$. Bei der Mutation wird mit Hilfe eines Zufallsgenerators ein Knotenpunkt ausgewählt und der Syntax entsprechend ausgetauscht, d.h., Operanden werden durch Operanden und Operatoren durch Operatoren ersetzt (Abbildung 2.2). Zusätzlich enthält das Programm eine modifizierte Version dieses Mutationstyps. Anstatt einen Operanden durch einen anderen Operanden auszutauschen, produziert das Programm einen neuen Teilbereich mit einer zu Beginn festgelegten maximal möglichen Knotenlänge und substituiert diesen entsprechend. Im Abschnitt der Sensitivitätsanalyse wird neben den traditionellen Evolutionsoperatoren ein als "Backpropagation" bezeichnetes Verfahren vorgestellt. Es stellt eine Annäherung an ein methodisches Verfahren aus dem Bereich der künstlichen Neuronalen Netze dar, in dem es die optimalen Gewichte der einzelnen Funktionsterme berechnet.

2.1 Aufbau und Ablauf des Genetischen Programms

Mit der Beschreibung der genetischen Operationen wurde bereits ein wesentlicher Teil des Genetischen Programms skizziert. Im weiteren soll der Aufbau und Ablauf des gesamten Programms näher erläutert werden. Helfen soll hierbei das Flußdiagramm 2.3, das den Ablauf schematisch zusammenfaßt. Neben anderen Modellparametern werden zu Beginn eine Funktions- und eine Terminalmenge definiert, die den Lösungsraum und damit einhergehend die Funktionsoperatoren, Variablen und Konstanten bestimmen, die in die

¹Die Klammersetzung erfolgt bei der Generierung der Anfangspopulation als auch bei einer weiter unten vorgestellten modifizierten Version eines Mutationstyps.

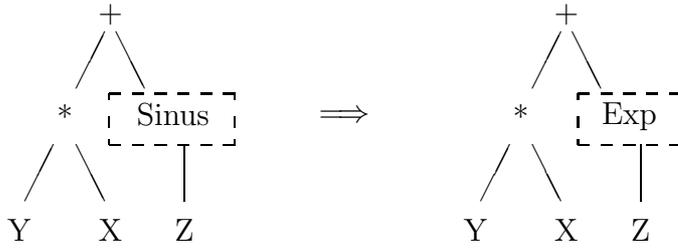


Abbildung 2.2: Mutation: Bei der Mutation muß darauf geachtet werden, daß die auszutauschenden Elemente mathematisch korrekt sind. So kann beispielsweise ein Plus-Operator nicht in eine Sinusfunktion umgewandelt werden.

mathematischen Gleichungen einfließen können. Die Wahrscheinlichkeit für die Operanden und Operatoren ausgewählt zu werden, ist jeweils gleichverteilt. In dieser Arbeit stehen folgende Funktions- und Terminalmengen zur Verfügung:

$$\text{Funktionsmenge} = \{+ \ - \ / \ * \ \sin \ \cos \ \ln \ \exp \ \sqrt{\cdot}\}$$

$$\text{Terminalmenge} = \{\text{Variable}, \mathbb{R} \in [0, 1], \mathbb{N} \in [1, 100]\}$$

Die Funktions- und Terminalmenge müssen sich nicht auf diese Komponenten beschränken, sondern sie hängen vielmehr vom Untersuchungsgegenstand ab². Für die kommenden Untersuchungen ist diese Wahl allerdings ausreichend. Da die Länge der mathematischen Gleichungen theoretisch keinen Restriktionen unterliegen, existieren unendlich viele Kombinationsmöglichkeiten mit einem unendlich großen Lösungsraum. Allerdings hat diese große Flexibilität keine positiven Auswirkungen auf die Ergebnisse. Ab einem bestimmten Zeitpunkt und einer bestimmten Größe der Gleichungen verursachen die genetischen Operationen ein sogenanntes “bloating”. Hierunter wird ein schnelles Anwachsen der Funktionslänge aufgrund redundanter mathematischer Terme bezeichnet. In der Literatur wird daher die Länge der mathematischen Gleichungen beschränkt. Diese Vorgehensweise wird hier ebenfalls angewendet, so daß die Funktionen eine Größe von maximal 100 Knoten annehmen können. Um den immens großen Lösungsraum sinnvoll zu durchsuchen, wird dem Genetischen Programm ein Suchalgorithmus zugrunde gelegt. Dieser Prozeß wurde bereits im Zusammenhang mit den Evolutionary Computation vorgestellt und ist das Zusammenspiel aus Darwins Selektionstheorie (“survival-of-the-fittest”) und der Anwendung genetischer Operationen, deren Funktionsweise der Genetik entspringt.

Das Genetische Programm erzeugt in der ersten Generation eine Anfangspopulation. Diese besteht aus einer vorher festgelegten Anzahl mathematischer Gleichungen, wobei jede Gleichung durch zufälliges Kombinieren der Operatoren und Operanden der Funktions- bzw. Terminalmenge sowie zufälliges Setzen von öffnenden und schließenden Klammern entsteht. Im Grundmodell weist die Anfangspopulation eine Größe von 100 Gleichungen

²Eine detaillierte Darstellung findet sich in Banzhaf et al. [5], S. 109 ff.

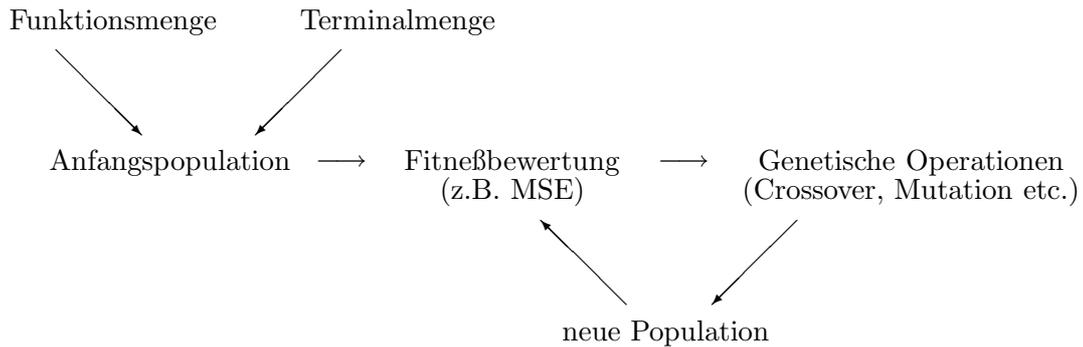


Abbildung 2.3: Flußdiagramm des Genetischen Programms: Die Auswahlwahrscheinlichkeit der Funktionen für die genetischen Operationen ist abhängig vom jeweiligen Fitneßwert.

auf. Bei der Festlegung der maximal möglichen Länge der Anfangsgleichungen kann die angegebene Knotenzahl K nicht exakt erreicht werden. Vielmehr bewegt sie sich im Bereich $2 * K - 1 \leq Knoten \leq 2 * (K - 1) + K$. Der Grund liegt im Aufbau des Computerprogramms: Die Generierung einer Gleichung der Anfangspopulation oder eines neuen Teilstrangs im Zusammenhang mit dem modifizierten Mutationstyp erfolgt in einer Computerschleife, die k -mal durchlaufen wird. Innerhalb dieser Schleife trifft das Programm zunächst auf die Terminalmenge und wählt mit Hilfe eines Zufallsgenerators einen Operanden aus. Im Anschluß an einen Operanden wählt das Programm einen Funktionsoperator. Da aber die Funktionsoperatoren \sin , \cos , \ln , \exp , $\sqrt{\cdot}$ nicht direkt auf eine Konstante oder einen Zeitreihenwert folgen dürfen, sondern einer der Operatoren $+$, $-$, $*$, $/$ zwischengeschaltet werden muß, wird die Funktionsmenge entsprechend aufgeteilt:

$$\text{Funktionsmenge 1} = \{+ \ - \ / \ *\}$$

$$\text{Funktionsmenge 2} = \{\sin \ \cos \ \ln \ \exp \ \sqrt{\cdot}\}$$

Das Programm wählt mit einer Wahrscheinlichkeit von 100% zunächst einen Operator aus Funktionsmenge 1 und mit einer Wahrscheinlichkeit von 50% einen Operanden aus der Funktionsmenge 2 aus. Für die Wahl eines Funktionsoperators innerhalb einer Funktionsmenge wird erneut ein Zufallsgenerator zu Hilfe genommen. Daher kann die Funktion für $k = 2$ bis zu vier Knoten aufweisen. Zusätzlich enthält das Programm die Bedingung, daß die Funktion mit einem Operanden enden muß, wenn die Schleife das letzte Mal durchlaufen wird.

Beim Erzeugen der mathematischen Gleichungen werden anhand eines Zufallsgenerators öffnende und schließende Klammern eingebaut. Wird eine öffnende Klammer gesetzt, so ändert sich die Wahrscheinlichkeit für das Setzen einer weiteren öffnenden Klammer auf $1/(1+\text{Anzahl öffnender Klammern})$, wobei das Setzen einer schließenden Klammer mit der entsprechenden Gegenwahrscheinlichkeit erfolgt.

Jede Gleichung stellt für sich eine mögliche Lösung dar. Um die Güte der zufällig erzeugten Gleichungen zu bewerten, wird für jede Gleichung ein Fitneßwert errechnet. Die

Evaluierung erfolgt üblicherweise anhand der mittleren quadratischen Fehlerabweichung (MSE):

$$MSE = \frac{1}{n} \sum_{i=1}^n (\hat{Y}_t - Y_t)^2 \quad (2.1)$$

mit n als Anzahl der Beobachtungswerte, Y_t als Zeitreihenwerte und \hat{Y}_t als geschätzte Zeitreihenwerte. Anschließend werden die Gleichungen gemäß ihres Fitneßwertes geordnet.

Durch Anwenden der genetischen Operationen wird in der nächsten Generation eine neue Population generiert. Die Wahrscheinlichkeit der Gleichungen für genetische Operationen ausgewählt zu werden, hängt im Grundmodell vom Fitneßrang ab. Je höher der Fitneßrang bzw. je kleiner die Fehlerabweichung, desto größer ist die Wahrscheinlichkeit für eine genetische Operation ausgewählt zu werden. Im Grundmodell werden die besten 25 Gleichungen für genetische Operationen herangezogen. Anschließend werden die Fitneßwerte der neuen Population berechnet und mit denen aus der Vorperiode verglichen. Die fittesten Gleichungen werden beibehalten und die restlichen verworfen, so daß an dieser Stelle Darwins Auslesetheorie berücksichtigt wird. Diese beiden Prozesse - das Generieren neuer Populationen anhand genetischer Operationen sowie die anschließende Selektion - wiederholen sich solange, bis entweder der Fitneßwert einer Lösung einen bestimmten Grenzwert oder das Programm eine bestimmte Anzahl an Iterationen erreicht hat. In diesem Fall hätte das Programm einen Lauf bzw. Run abgeschlossen.

Die nachfolgende Tabelle 2.1 listet die Parameterwerte für das Grundmodell auf. Ausgehend von dieser Konfiguration werden im Abschnitt der Sensitivitätsanalyse einzelne Parameter variiert, um ihren Einfluß auf die Fitneßentwicklung zu überprüfen.

Anfangsparameter des GP GAUSS Grundmodells		
Anzahl der Läufe	=	100
Anzahl der Generationen pro Lauf	=	200
Populationsgröße	=	100
Fitneßkriterium	=	MSE
maximale Länge der Funktionen der Anfangspopulation	=	5 Knoten
maximale Länge der Funktionen im weiteren Verlauf	=	100 Knoten
Anzahl neuer Gleichungen durch Crossoveroperationen	=	60
Anzahl neuer Gleichungen durch Mutationsoperationen	=	40
Typ A:	=	20
Typ B:	=	20
maximale Länge der Funktionsteile bei Mutationstyp B	=	5 Knoten
Anzahl der besten Funktionen, die für genetische Operationen herangezogen werden	=	25

Tabelle 2.1: Mutationstyp A: Operand wird durch einen Operand und Operator durch einen Operator ersetzt. Mutationstyp B: Austausch eines Operanden durch einen neu generierten Funktionsteil.

2.2 Vorteile des Genetischen Programms

Nachdem das Genetische Programm detailliert erläutert wurde, sollen die Vorteile des Programms, die sich vor allem im Hinblick auf Finanzmarktprognosemodelle ergeben, dargelegt werden. Der Vergleich des Genetischen Programms mit den hier vorgestellten Forschungszweigen beginnt erneut mit den künstlichen neuronalen Netzen.

Ein häufig erwähnter Kritikpunkt neuronaler Netze stellt ihre Black-Box Eigenschaft dar, deren nicht beobachtbarer Optimierungsprozeß keine ökonomisch interpretierbare Regressionsgleichung erzeugt. Somit könnten erfolgreiche neuronale Prognosemodelle für den Finanzmarkt zwar belegen, daß die Zeitreihen nichtlineare Elemente besitzen, aber es wäre nicht möglich, deren konkrete Ausgestaltung offenzulegen. Dieser Nachteil existiert beim Genetischen Programm nicht. Zu jedem Zeitpunkt kann sich der Anwender sämtliche Lösungskandidaten und bei Bedarf die Gestalt und Auswirkungen genetischer Operationen anzeigen lassen. Allerdings stellt diese Transparenz nur bedingt einen Vorteil dar, da eine ökonomische Interpretation aufgrund der häufig vorliegenden Komplexität der Gleichungen nur eingeschränkt möglich ist.

Wenig sinnvoll erscheint hingegen ein Vergleich zwischen Fuzzy Systemen und dem Genetischen Programm, da die Fuzzy Systeme nicht primär zur Schätzung von Zeitreihen konzipiert wurden. Vielmehr stehen Fuzzy Systeme in Bezug auf die hier angeführten Forschungsbereiche in einer komplementären und nicht in einer konkurrierenden Beziehung, wie zahlreiche Publikationen belegen.

Eine ähnliche Situation ergibt sich beim Forschungsbereich Evolutionary Programming, deren häufigste Anwendung im Bereich der Mustererkennung zu finden ist. Den interessantesten Vergleich bieten die Genetischen Algorithmen. Sämtliche Nachteile, die sich mit der Kodierung des Lösungsraumes einstellen, entfallen beim Genetischen Programm. Beim Genetischen Programm existiert keine problemspezifische, zeitaufwendige Kodierung und Zuordnung von Eigenschaften, die bei unzureichender Form die Ergebnisgüte negativ beeinflussen können. Des weiteren wurde die Restriktion der fixen Länge der Lösungskandidaten eliminiert, so daß die Restriktion des Lösungsraums aufgehoben wird und ein theoretisch unendlich großer Lösungsraum durchsucht werden kann. Im Hinblick auf die genetischen Operationen gestaltet sich die Umsetzung bei den Genetischen Algorithmen zwar erheblich einfacher, aber aufgrund der diskreten Änderungen im Gegensatz zum Genetischen Programm besitzen sie zugleich den weiteren Nachteil einer nur annäherungsweise optimalen Lösung.

Insgesamt bleibt festzuhalten, daß jeder Forschungsbereich seine Stärken und Schwächen aufweist. Richtet sich das Ziel der Untersuchung aber auf die Schätzung und Prognose von Finanzmarktzeitreihen, so bietet das Genetische Programm wegen des Wegfalls der den anderen Forschungsbereichen inhärenten Nachteile zumindest im Bereich der hier vorgestellten computerbasierten Suchalgorithmen einige Vorteile.

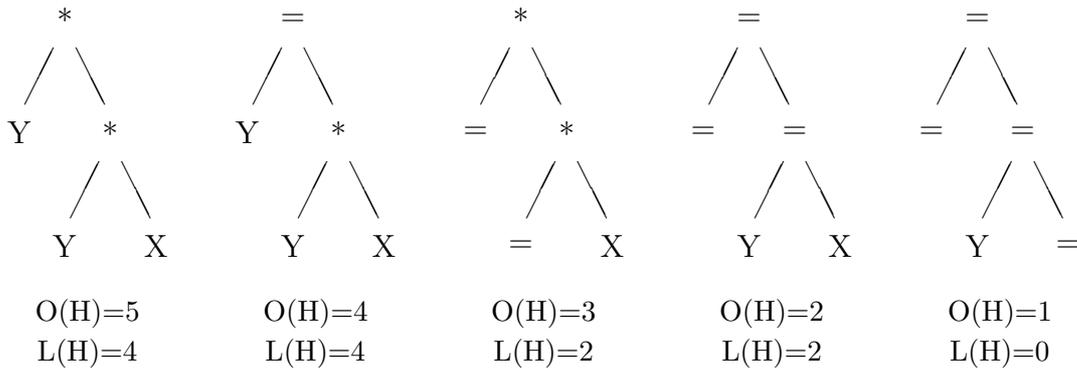


Abbildung 2.4: Verschiedene Schemata, mit denen der Lösungskandidat $Y*(Y*X)$ übereinstimmt. Die Knotenanzahl der Schemata beträgt $N(H) = 5$. $O(H)$: Anzahl der Symbole unterschiedlich von '='; $L(H)$: Anzahl der Verbindungen für minimales Fragment ohne '='.

2.3 Schema Theorem

Nachdem im Abschnitt 2.1 der Aufbau und Ablauf Genetischer Programme vorgestellt worden ist, soll ein Schema Theorem nach Poli & Langdon [105] (1998) vorgestellt werden, das die Dynamiken der Populationsentwicklung bzw. die Entwicklung der Strukturen innerhalb der einzelnen Lösungsgleichungen mathematisch beschreibt³. Holland [64] entwickelte 1975 einen derartigen mathematischen Ansatz bereits für Genetische Algorithmen. Es basiert auf dem Konzept von Schemata, die eine Art Schablone darstellen und mit Teilen der binären Strängen einzelner Lösungskandidaten übereinstimmen und somit eine Teilmenge des gesamten Lösungsraumes beschreiben. Um diese Aufgabe besser erfüllen zu können, verfügen die Schemata zusätzlich über ein sogenanntes „Don't Care“ Symbol, das Holland durch ein „*“ Zeichen kennzeichnet. Zum Beispiel beinhaltet das Schema 0^*10 die Menge der beiden Lösungskandidaten 0010 und 0110 . Umgekehrt kann ein binärer Strang exakt 2^n verschiedene Schemata repräsentieren, wobei n die Anzahl der Ziffern eines Binärstrangs definiert. Wenn $m(s, t)$ die Anzahl der Übereinstimmungen mit dem Schema s in der Population zum Zeitpunkt t kennzeichnet, dann beschreibt das Schema Theorem den erwarteten Wert $m(s, t + 1)$ in Abhängigkeit von $m(s, t)$, der Größe der Population, den Evolutionsoperatoren sowie anderen Eigenschaften des Schemata. Im Hinblick auf Genetisches Programmieren existieren mehrere Theoreme, die den Ansatz von Holland zu übertragen versuchen⁴. Allerdings ist die Erstellung eines Schema Theorems für GP aufgrund der variablen Länge der Lösungskandidaten und der unterschiedlichen Positionen der ausgetauschten Teilbereiche innerhalb der Baumstrukturen wesentlich komplexer. Insgesamt kann man die GP Schema Theoreme in „non-rooted“ und „rooted“ basierte Ansätze einteilen. Zu den „non-rooted“ basierten Ansätzen zählen die Theoreme von Koza (1992) [75], O'Reilly (1995) [99] und Whigham (1996) [129], die sich dadurch auszeichnen, daß die Baumstruktur eines Lösungskandidaten an beliebigen Positionen mit dem Schema abgeglichen wird und folglich ein Schema innerhalb eines Lösungskandidaten

³Der Abschnitt über das Schema Theorem orientiert sich an dem Artikel von Poli & Langdon.

⁴Siehe auch Banzhaf et al. [5] und Poli & Langdon [105].

vielfach auftreten kann. Im Gegensatz dazu setzen die „rooted“ basierten Theoreme stets am obersten Knotenpunkt der Baumstruktur an, so daß ein Schema höchstens einmal auftauchen kann. Dieser Ansatz wird bei Rosca (1997) [111] und Poli & Langdon [105] angewendet und bietet den Vorteil, Wahrscheinlichkeiten leichter berechnen zu können.

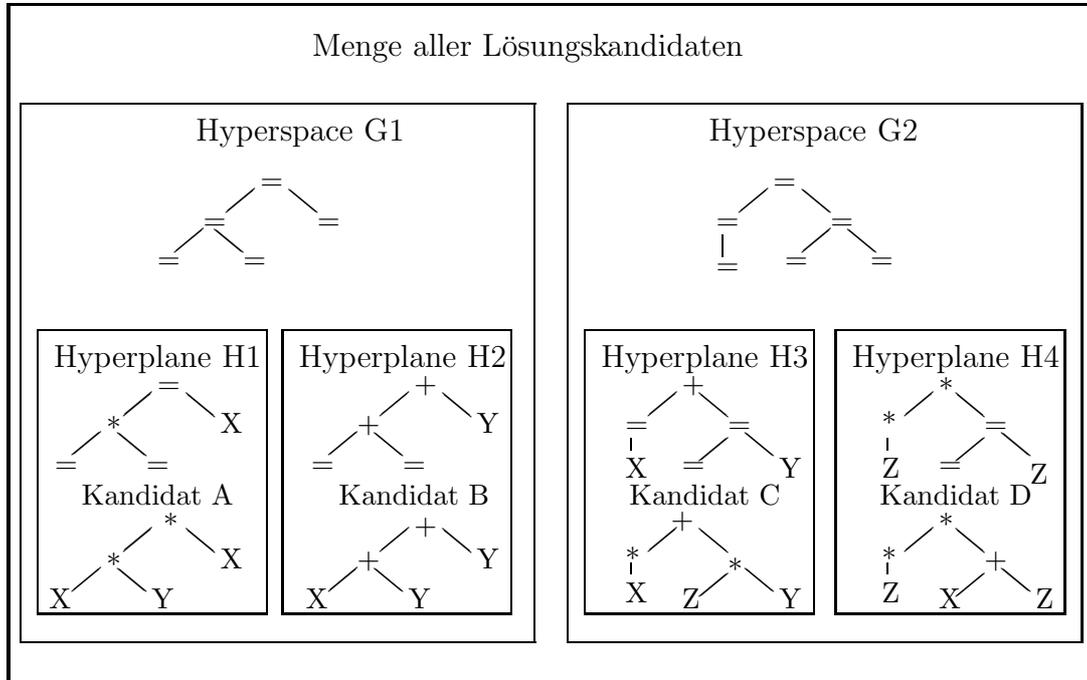


Abbildung 2.5: Schema H nach Poli & Langdon: Alle Lösungskandidaten werden dahingehend überprüft, ob ihre Baumstrukturen das Schema H aufweisen. Die Anzahl der Verbindungen $L(H)$ sind durch die Zahlen 1 bis 3 dargestellt.

2.3.1 GP Schema Theorem nach Poli & Langdon

Ähnlich wie beim Schema Theorem von Holland verwenden auch Poli & Langdon zum Erstellen der Schemata ein „Don't Care“ Symbol, das sie durch ein „=“ Zeichen darstellen. Dabei kann das „Don't Care“ Symbol eines der Elemente aus der Funktions- oder Terminalmenge annehmen. Poli & Langdon definieren ein Schema folgendermaßen:

1. **Definition: GP Schema.** Ein GP Schema ist die gesamte oder ein Teilbereich der Baumstruktur eines Lösungskandidaten bestehend aus den Operatoren und Operanden der Funktions- und Terminalmenge. Das Element „=“ stellt eine polymorphe Funktion dar, die jeden Eintrag der Funktions- oder Terminalmenge annehmen kann.
2. **Definition: Ordnung.** Die Ordnung $O(H)$ eines Schemas H ist die Anzahl der Symbole, die keine „Don't Care“ Symbole „=“ darstellen.

3. **Definition: Länge.** Die Länge $N(H)$ eines Schemas H ist die Anzahl der Knoten im Schema H .
4. **Definition: Definierte Länge.** Die definierte Länge $L(H)$ eines Schemas H ist die Anzahl aller Verbindungen, die man benötigt, um das minimale Fragment von H ohne die „Don't Care“ Symbole zu erzeugen.
5. **Definition: Hyperspaces und Hyperplanes.** Ein Schema H ist ein Hyperspace (Hyperplane), wenn $O(H) = 0$ ($O(H) > 0$) erfüllt ist. Durch Ersetzen sämtlicher Symbole, die **keine** „Don't Care“ Symbole darstellen, durch „Don't Care“ Symbole enthält man den entsprechenden Hyperspace des Hyperplanes H .

In Abbildung 2.4 sind einige Schemata mit den dazugehörigen Ordnungen, Längen und definierten Längen dargestellt. Der Zusammenhang zwischen Hyperspace, Hyperplane, Schema und Lösungskandidat wird in Abbildung 2.5 graphisch erläutert. Ein Schema mit einer niedrigen Ordnung und großer Länge kann eine sehr große Anzahl an Lösungskandidaten beinhalten. Wenn beispielsweise die Funktions- und Terminalmenge aus nur zwei Operatoren und Operanden besteht, repräsentiert ein Schema mit der Ordnung $O(H) = 5$ und Länge $N(H) = 30$ insgesamt $2^{N-O} = 33.554.432$ Lösungskandidaten.

One-Point-Crossover

Poli & Langdon führen einen neuen mit One-Point-Crossover bezeichneten Operator ein, der neben anderen Aspekten vor allem die Berechnung der Wahrscheinlichkeit für die Zerstörung von GP Schemata ermöglicht. Der Vorgang der One-Point-Crossover Operationen setzt sich aus den drei folgenden Phasen zusammen und ist in Abbildung 2.6 graphisch dargestellt:

1. **Abgleichung.** In einem ersten Schritt werden die Baumstrukturen zweier Lösungskandidaten betrachtet und die Teilbereiche mit der gleichen Gestalt und Knotenanzahl beginnend vom obersten Knotenpunkt identifiziert. Die übereinstimmenden Verbindungen werden gespeichert.
2. **Crossover Point Selektion.** Innerhalb der gemeinsamen Strukturen wird anhand eines Zufallsgenerators ein Knotenpunkt gewählt, der für beide Lösungskandidaten gilt und an dem die Teilstränge abgeschnitten werden.
3. **Crossover.** Die Teilstränge werden wie beim Standardverfahren eines Crossovers ausgetauscht.

Im Zusammenhang mit den One-Point-Crossover Operationen sollen zwei weitere von Poli & Langdon erwähnte Aspekte angebracht werden, die sich im Unterschied zum Standardverfahren einstellen.

Konvergenzeigenschaft Durch das Verwenden von One-Point-Crossover Operationen und bei Abwesenheit von Mutationsoperationen konvergieren die Strukturen innerhalb

einer Population. Somit bleibt die Wahrscheinlichkeit einen Crossover Point im unteren Bereich einer Baumstruktur auszuwählen solange relativ klein, bis sich die oberen Bereiche der Baumstrukturen einer Population bis zu einem gewissen Grad angeglichen haben. Anders ausgedrückt bedeutet dies, daß die Wahrscheinlichkeit, kleinere Fragmente auszutauschen, im Zeitablauf steigt. Eine wichtige Rolle nehmen die Mutationsoperationen ein, die eine bestimmte Diversifikation in den Baumstrukturen aufrechterhalten und das Programm vor einer vorzeitigen Konvergenz bewahren.

Tiefe der Baumstruktur Weiterhin ist eine Folge der One-Point-Crossover Operationen, daß die Tiefe der Baumstruktur bzw. die Knotenanzahl der neu geschaffenen gegenüber den ursprünglichen Lösungskandidaten nicht anwächst. Dies bedeutet, daß die Performance für Genetische Programme mit One-Point-Crossover und Point Mutationsoperationen sich auf einen Lösungsraum beschränkt, der durch die Anfangspopulation aufgespannt wird.

Schema Theorem

Mit den oben gegebenen Definitionen läßt sich ein Theorem herleiten, das die Änderung der Anzahl an Lösungskandidaten, die einem bestimmten Schema zugeordnet werden können, innerhalb eines Generationsschritts von t nach $t + 1$ beschreibt. Dabei hängt die Verbreitung eines Schemas von den Faktoren des Selektionsprozesses, den Crossover- und Mutationsoperationen ab und basiert auf den Wahrscheinlichkeiten, daß ein Schema H nicht durch genetische Operationen zerstört wird:

$$E[m(H, t + 1)] = M * Pr\{h \in H\} * (1 - Pr\{D_m(H)\}) * (1 - p_c * Pr\{D_c(H)\}) \quad (2.2)$$

Demnach berechnet sich die erwartete Anzahl $E[m(H, t + 1)]$ der Lösungskandidaten, die mit Schema H in Generation $t + 1$ übereinstimmen, durch die Multiplikation der Populationsgröße M und den Wahrscheinlichkeiten, daß ein Kandidat h aus der Schemamenge H ausgewählt wurde ($Pr\{h \in H\}$) und das Schema H nicht durch eine Mutationsoperation ($1 - Pr\{D_m(H)\}$) oder durch eine Crossoveroperation ($1 - p_c * Pr\{D_c(H)\}$) zerstört wird. Dabei bezeichnen die Parameter D_m und D_c die Ereignisse einer Zerstörung des Schemas H durch eine Mutation oder einer Crossoveroperation. Der Term p_c gibt die Crossover-Wahrscheinlichkeit an.

Bei einem fitneßabhängigen Selektionsverfahren läßt sich die Selektionswahrscheinlichkeit $Pr\{h \in H\}$ formulieren als

$$Pr\{h \in H\} = \frac{m(H, t) * f(H, t)}{M \bar{f}(t)} \quad (2.3)$$

wobei $f(H, t)$ die durchschnittliche Fitneß der Lösungskandidaten der Schemamenge H und $\bar{f}(t)$ den durchschnittlichen Fitneßwert der gesamten Population ausdrücken.

Da ein Schema H eine Mutation nur überlebt, wenn ausschließlich die Knoten eines

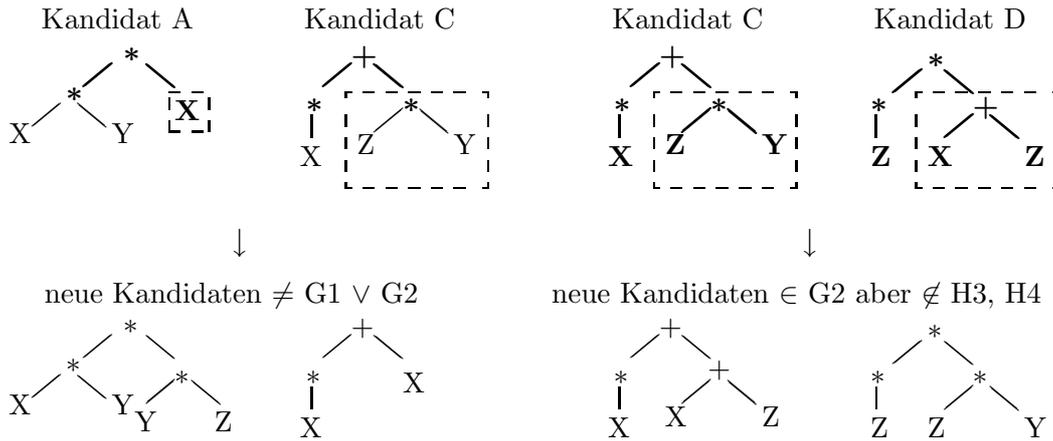


Abbildung 2.6: Poli & Langdon: zwei Beispiele für die Zerstörung eines Schemas durch One-Point-Crossover. Die fettgedruckten Bereiche heben die übereinstimmenden Verbindungen hervor (Ableichungsvorgang). Der gestrichelte Rahmen kennzeichnet die auszutauschenden Teilbereiche (Crossover Point Selektionsvorgang).

Lösungskandidaten betroffen sind, die beim Schema H einem „Don't Care“ Symbol entsprechen, beträgt die Wahrscheinlichkeit, daß ein Schema H durch eine Mutation zerstört wird

$$Pr\{D_m(H)\} = 1 - (1 - p_m)^{O(H)} \quad (2.4)$$

mit p_m als Mutationswahrscheinlichkeit. Schemata mit einer größeren Ordnung $O(H)$ werden demnach häufiger zerstört.

Die Formulierung der Wahrscheinlichkeiten für die One-Point-Crossover Operationen gestaltet sich komplizierter. Es existieren zwei Ereignisse, die ein Schema H zerstören können⁵. Der erste Fall⁶ ($D_{c1}(H)$) beschreibt einen Crossovervorgang zwischen zwei Lösungskandidaten h und \hat{h} , die unterschiedliche Schemata aufweisen. Wenn $G(H)$ der Hyperspace von Schema H ist, dann berechnet sich die Wahrscheinlichkeit für D_{c1} als

$$Pr\{D_{c1}(H)\} = Pr\{D_c(H)|\hat{h} \notin G(H)\} * Pr\{\hat{h} \notin G(H)\}. \quad (2.5)$$

Der Term $Pr\{\hat{h} \notin G(H)\}$ beschreibt die Wahrscheinlichkeit, einen Lösungskandidaten auszuwählen, der nicht das Schema H aufweist. Unter der Annahme einer fitneßproportionalen Selektion kann diese Wahrscheinlichkeit formuliert werden als

$$Pr\{\hat{h} \notin G(H)\} = 1 - \frac{m(G(H), t) * f(G(H), t)}{M\bar{f}(t)} \quad (2.6)$$

Dagegen läßt sich der Term $Pr\{D_c(H)|\hat{h} \notin G(H)\}$ nur schwer berechnen, da einige Crossoveroperationen zwischen Kandidaten mit verschiedenen Schemata neue Lösungen generieren können, die dennoch mit dem Schema H übereinstimmen. Im weiteren Verlauf wird

⁵Siehe auch Abbildung 2.6

⁶Abbildung 2.6: Crossovervorgang zwischen Kandidat A und C.

der Term mit $p_{diff}(t)$ abgekürzt.

Beim zweiten Fall⁷ ($D_{c2}(H)$) handelt es sich um Crossoveroperationen zwischen zwei Lösungskandidaten h und \hat{h} , wobei \hat{h} zwar die gleiche Struktur wie das Schema H besitzt ($\hat{h} \in G(H)$), aber nicht mit diesem übereinstimmt ($\hat{h} \notin H$). Ein solches Ereignis kann durch den Term $D_{c2}(H) = \{D_c(H), \hat{h} \notin H, \hat{h} \in G(H)\}$ ausgedrückt werden mit der entsprechenden Wahrscheinlichkeit:

$$Pr\{D_{c2}(H)\} = Pr\{D_c(H)|\hat{h} \notin H, \hat{h} \in G(H)\} * Pr\{\hat{h} \notin H, \hat{h} \in G(H)\} \quad (2.7)$$

Wenn die Ereignisse $\{\hat{h} \in H\}$ eine Teilmenge der Ereignisse $\{\hat{h} \in G(H)\}$ mit $H \subset G(H)$ sind, dann läßt sich der zweite Term der Gleichung 2.7 ausdrücken als

$$Pr\{\hat{h} \notin H, \hat{h} \in G(H)\} = Pr\{\hat{h} \in G(H)\} - Pr\{\hat{h} \in H\} \quad (2.8)$$

mit

$$Pr\{\hat{h} \in G(H)\} = \frac{m(G(H), t) * f(G(H), t)}{M * \bar{f}(t)} \quad (2.9)$$

und

$$Pr\{\hat{h} \in H\} = \frac{m(H, t) * f(H, t)}{M * \bar{f}(t)}. \quad (2.10)$$

Damit Schema H zerstört wird, muß der Crossover Point zwischen den Knoten liegen, die nicht zum „Don't Care“ Symbol gehören. Die Wahrscheinlichkeit hierfür beträgt $\frac{L(H)}{N(H)-1}$. Zusammen mit den bisherigen Gleichungen und der Ungleichung

$$Pr\{D_c(H)|\hat{h} \notin H, \hat{h} \in G(H)\} \leq Pr\{B(H)\} \quad (2.11)$$

erhält man:

$$Pr\{D_{c2}(H)\} \leq \frac{L(H)}{N(H)-1} * \frac{m(G(H), t) * f(G(H), t) - m(H, t) * f(H, t)}{M * \bar{f}(t)} \quad (2.12)$$

Die Wahrscheinlichkeit für $Pr\{D_c(H)\}$ ergibt sich aus der Summe $Pr\{D_{c1}(H)\}$ und $Pr\{D_{c2}(H)\}$. Durch Einsetzen der einzelnen Wahrscheinlichkeiten in die Gleichung 2.2 gelangt man schließlich zum GP Schema Theorem nach Poli & Langdon:

$$E[m(H, t + 1)] \geq m(H, t) * \frac{f(H, t)}{M * \bar{f}(t)} * (1 - p_m)^{O(H)} * \left\{ 1 - p_c * \left[p_{diff}(t) * \left(1 - \frac{m(G(H), t) * f(G(H), t)}{M \bar{f}(t)} \right) + \frac{L(H)}{N(H)-1} * \frac{m(G(H), t) * f(G(H), t) - m(H, t) * f(H, t)}{M * \bar{f}(t)} \right] \right\} \quad (2.13)$$

⁷Abbildung 2.6: Crossovervorgang zwischen Kandidat C und D.

Analyse des GP Schema Theorems

In ihrem Artikel diskutieren Poli & Langdon, welche Gewichtung die einzelnen mathematischen Terme der Gleichung 2.13 im Zeitablauf theoretisch einnehmen müßten. Demnach läßt sich die dynamische Entwicklung der Schemata in zwei Phasen einteilen. Die Autoren behaupten, daß in den ersten Generationen (Phase I) eines GP Laufs die Wahrscheinlichkeit für die Zerstörung eines Schemas aufgrund der heterogenen Struktur der Anfangspopulation sehr groß ist. Daher konkurrieren zunächst die verschiedenen Hyperspaces $G(H)$ miteinander mit der Folge, daß als Erstes die Gestalt und Größe der Schemata im Vordergrund steht. Dabei ist ein Hyperspace um so erfolgreicher, je mehr Lösungskandidaten es anhäuft und je besser seine durchschnittliche Fitneß ist. In einer zweiten Phase beginnen die Hyperplanes bzw. Schemata bei den verbliebenen Hyperspaces untereinander zu konkurrieren, wobei hier die definierte Länge und die Fitneß des Schemas entscheidend sind.

2.4 Bloating

Eine typische Eigenschaft, die einem Genetischen Programm zugrunde liegt, ist ein starkes Anwachsen der Lösungskandidaten in Bezug auf deren Knotenlänge. Das Wachstum wird lediglich durch die vorher festgelegte maximale Knotenlänge eines Individuums begrenzt. Diese Eigenschaft wird in der Literatur häufig mit dem Begriff 'bloating' umschrieben⁸ und wurde bereits in der Anfangszeit des Genetischen Programmierens durch Takett [121] (1994) offengelegt. Im gleichen Jahr fand der Autor Angeline [2] heraus, daß die Lösungskandidaten im Zeitablauf Bereiche aufweisen, die keinen Einfluß bei der Berechnung der Fitneß ausüben. Angeline bezeichnet diese überflüssigen Segmente als 'Introns' und orientiert sich damit an dem biologischen Fachausdruck für überflüssiges DNA Material, das keine Wirkung auf den Phänotyp eines Individuums hat. Die in der Folgezeit entstandenen Studien sahen in den Introns die Ursache für das Bloating (Nordin & Banzhaf [97], Nordin et al.[98], Soule et al. [116], Rosca [111]), und daß sie eine persistente und problematische Rolle darstellen. So zeigen einige Studien, daß im ersten Abschnitt eines Laufs etwa 40-60 % eines Individuums aus überflüssigem Material besteht und daß sich diese Introns im weiteren Verlauf exponentiell fortpflanzen. Die Auswirkungen des Bloatings werden in der Literatur insgesamt als negativ angesehen, wengleich manche Autoren dem Bloating auch nützliche Eigenschaften zuschreiben. Eindeutig negativ zu bewerten ist die Tatsache, daß aufgrund des exponentiellen Wachstums und aufgrund der Begrenzung der Knotenanzahl die Fitneßentwicklung zu stagnieren beginnt, was sich auch in einer hyperbolischen Fitneßentwicklung ausdrückt. Auf der anderen Seite betont Angeline, daß gerade dieses zusätzliche genetische Material für die Fitneßentscheidung wichtig ist.

⁸Siehe auch Banzhaf et al. [5], S.181 ff.

2.5 Residuenanalyse

Bevor GP GAUSS im nächsten Abschnitt auf dynamische Systeme angewendet wird, sollen kurz die beiden Testverfahren auf Normalverteilung und Autokorrelation vorgestellt werden, anhand derer die Schätzresiduen ϵ_t überprüft werden. Für diesen Vorgang werden die Fehler ϵ_t zunächst durch Division der Standardabweichung σ_ϵ normalisiert: $u_t = \epsilon_t/\sigma_\epsilon$.

2.5.1 Test auf Normalverteilung

Ein häufig verwendeter Test auf Normalverteilung ist der Jarque-Bera-Test [69]. Er basiert auf Schätzungen des dritten und vierten Moments einer Verteilung und ist asymptotisch χ^2 -verteilt mit zwei Freiheitsgraden:

$$JB = (n - p)(S^2/6 + (K - 3)^2/24) \quad (2.14)$$

mit n als die Anzahl der Beobachtungswerte und p als die Anzahl der Parameter in einem Regressionsmodell. Die Schiefe S als drittes und die Kurtosis K als viertes Moment werden folgendermaßen definiert:

$$S(X) = E(X - \mu)^3/\sigma_u^3 \quad (2.15)$$

$$K(X) = E(X - \mu)^4/\sigma_u^4 \quad (2.16)$$

mit X als Zeitreihenwerte und μ als arithmetisches Mittel der Zeitreihenwerte. Die Nullhypothese normalverteilter Residuen wird abgelehnt, falls die Jarque-Bera Teststatistik den kritischen Wert der $\chi_\alpha^2(2)$ -Verteilung bei einer Irrtumswahrscheinlichkeit von α -Prozent überschreitet. Darüber hinaus können die Schiefe- und Kurtosiskoeffizienten interessante Informationen offenlegen. Kurtosiswerte von $K > 3$ signalisieren eine leptokurtotische Verteilung der Fehlerterme, was bedeutet, daß das Auftreten von Prognosefehlern im Vergleich zur Normalverteilung im Bereich der Schultern eine geringere, um den arithmetischen Mittelwert und in den Extrembereichen eine größere Wahrscheinlichkeit besitzt. Leptokurtotisches Verhalten tritt häufig im Zusammenhang mit Außreißerwerten bei den Prognosen auf. Weichen dagegen der Median und der arithmetische Mittelwert voneinander ab, so treten bei einer rechtsschiefen (linksschiefen) Verteilung zu hohe (niedrige) Prognosen häufiger auf als zu niedrige (hohe) Prognosen. Dieser Aspekt könnte vor allem bei den späteren Anwendungen genetisch entwickelter Prognosemodelle auf Finanzmarktzeitreihen interessant sein.

2.5.2 Test auf Autokorrelation

Der Test auf autokorrelierte Fehlerterme soll anhand der Box-Ljung Methode vorgenommen werden. Die Box-Ljung [80] Teststatistik berechnet sich als:

$$BL = n(n + 2) \sum_{k=1}^m \frac{\rho_k^2}{(n - k)} \quad (2.17)$$

mit k als Index zeitverzögerter Werte. Der Box-Ljung Test ist ebenfalls χ^2 -verteilt mit insgesamt m Freiheitsgraden. Der Autokorrelationskoeffizient ρ_k^2 wird durch die Relation zwischen Kovarianz und Varianz ermittelt:

$$\rho_k^2 = \frac{\text{cov}(X_t, X_{t+k})}{\text{var}(X_t)} = \frac{\sum (X_t - \mu(X))(X_{t+k} - \mu(X))}{\sum (X_t - \mu(X))^2} \quad (2.18)$$

Sind die Fehlerterme voneinander nicht unabhängig, deutet dies auf das Fehlen einer erklärenden Variablen in der Regressionsgleichung und somit auf eine Fehlspezifikation hin. Die Nicht-Erfüllung normalverteilter Residuen bei Prognosemodellen wird im Allgemeinen im Vergleich zu autokorrelierten Residuen als weniger gravierend angesehen.

2.6 Symbolische Regressionen und Vergleich des GP GAUSS mit anderen Programmen

Um herauszufinden wie gut das Genetische Programm funktioniert, soll es unterschiedlich komplexe dynamische Strukturen anhand von In-Sample Schätzungen approximieren. Die ausgewählten Zeitreihen stammen aus den beiden Aufsätzen von Kaboudan [72] und Kaboudan & Vance [73]. Zunächst soll GP GAUSS mit den Ergebnissen aus dem Artikel von Kaboudan & Vance verglichen werden, die in ihrem Artikel drei Genetische Programme mit unterschiedlichen Quellcodes miteinander vergleichen. Es handelt sich dabei zum einen um das von Fraser konzipierte Programm GPCPP (1994) sowie um das von Singleton geschriebene GPQuick (1995) und Horner's Vienna University GP Kernel (1996). Einige Zeitreihen aus diesen beiden Artikeln wurden von Kaboudan freundlicherweise bereitgestellt, so daß ein direkter Vergleich der verschiedenen Genetischen Programme ermöglicht wird. Die in diesem Abschnitt angeführten Untersuchungen konzentrieren sich auf die zur Verfügung gestellten Zeitreihen, so daß nicht alle dynamischen Prozesse aus Gründen mangelnder Vergleichbarkeit berücksichtigt wurden. Ein weiteres Genetisches Programm, das sogenannte Time Series Genetic Programming (TSGP), stellt eine Weiterentwicklung des GPQuick von Kaboudan dar. Die Software ist im Internet frei verfügbar, so daß die Zeitreihen auch auf dieses Programm angewendet werden⁹. Folgende dynamische Prozesse sollen geschätzt werden:

Lineare Funktion OZ (Ozaki [100], 1985):

$$Y_t = 1.8708 * Y_{t-1} - Y_{t-2} \quad (2.19)$$

Nichtlineare chaotische Funktion - Logistic Map LG (May [87], 1976):

$$Y_t = 4 * Y_{t-1}(1 - Y_{t-1}) \quad (2.20)$$

Nichtlinear chaotische Funktion - Henon Map HN (Henon [61], 1976):

$$Y_t = 0.3 * Y_{t-2} + 1 - 1.4 * Y_{t-1}^2 \quad (2.21)$$

⁹Die Software ist auf der folgenden Internetseite erhältlich: http://newton.uor.edu/facultyfolder/mahmoud_kaboudan

Nichtlineare trigonometrische Funktion TF:

$$Y_t = 3.9 * \sin(Y_{t-1}) - 0.85 * \cos(Y_{t-2}) \quad (2.22)$$

Kaboudan & Vance verwendeten für die untersuchten Programme eine Populationsgröße von 100 Gleichungen und 100 Läufen pro System. Zur Anzahl der Generationen gibt es leider keine Angaben. Neben diesen Parametern wurde daher für TSGP und GP GAUSS eine Generationenanzahl von 1000 festgelegt. Die Güte der mit den Genetischen Programmen

DGP:	GPCPP	GPQuick	GPV	TSGP	GP GAUSS
OZ: R^2	0.997	0.998	0.999	0.997	0.999
St.Dv.	0.002	0.001	0.000	0.002	0.000
LG: R^2	0.987	0.995	0.868	0.942	0.997
St.Dv.	0.008	0.004	0.095	0.039	0.002
HN: R^2	0.933	0.957	0.879	0.891	0.989
St.Dv.	0.030	0.018	0.021	0.031	0.007
TF: R^2	0.962	0.986	0.897	0.961	0.987
St.Dv.	0.016	0.011	0.090	0.037	0.011

Tabelle 2.2: Vergleich Genetischer Programme anhand des Bestimmtheitsmaßes. Die Ergebnisse der Spalten 2-4 wurden aus Kaboudan & Vance [73] entnommen, während die Resultate der Spalte 5 mit Hilfe der TSGP Software generiert wurden.

entwickelten Modelle ermittelt sich anhand des Bestimmtheitsmaßes R^2 , wobei die Selektion nach wie vor anhand der durchschnittlichen quadratischen Fehlerabweichung erfolgte. Zur Berechnung des Bestimmtheitsmaßes wurden die Daten von 100 Läufen gesammelt und der Durchschnitt der besten 25% ermittelt. Die Spalten 2 – 4 der Tabelle 2.2 zeigen die Ergebnisse der von Kaboudan & Vance untersuchten Programme. Man erkennt, daß GPQuick die Funktionen am besten approximiert. Betrachtet man die Ergebnisse des von Kaboudan weiterentwickelten GPQuick Computerprogramms TSGP, so fallen die Ergebnisse eher ernüchternd aus. Anstelle einer erwarteten Verbesserung gegenüber GPQuick weist es sogar schlechtere Resultate als GPCPP auf. Auch das mögliche Argument zufallsbedingter Ergebnisse überzeugt aufgrund des untersuchten Datenumfanges nur wenig. In diesem Umfeld stellt sich nun die Frage, wie gut das vorliegende Genetische Programm im Vergleich zu den anderen abschneidet. In Spalte sechs sind die Ergebnisse des GP GAUSS dargestellt. Demzufolge ist das vorliegende Programm in der Lage, die dynamischen Strukturen ohne Ausnahme besser zu schätzen. Allerdings benötigt GP GAUSS eine wesentlich größere Rechenzeit als TSGP und daher vermutlich auch als GPQuick. Ein möglicher Grund für deren Schnelligkeit könnte darin bestehen, daß die Computersprache C++ die Basis der Programme bildet.

Anhand der Abbildung 2.7 auf Seite 35 soll das Ergebnis einer der symbolischen Regressionen veranschaulicht werden. Die Graphik bildet sowohl die Originalzeitreihe der nichtlinearen trigonometrischen Funktion TF als auch die Zeitreihe der besten Schätzgleichung, die das GP GAUSS Programm gefunden hat, ab. Das Genetische Programm approximiert

den dynamischen Prozeß allerdings so gut, daß die Fehlerabweichung kaum sichtbar ist. Die entsprechende von GP GAUSS entwickelte Gleichung sieht folgendermaßen aus:

$$\begin{aligned}
TF = & \left(\left(\left(\left(\sin(Y_t[v - 0001]) + (\sin(\sin(Y_t[v - 0001]))) + (\sin(Y_t[v - 0001])) \right) \right) \right) \right) + (\sin(\sin(\sin(Y_t[v - 0001]))) + (\cos(\sin(\sin(Y_t[v - 0001]) + (\ln1((0.30)))) + ((\cos(wzl((\sin(Y_t[v - 0002])))))))) + (\cos((Y_t[v - 0002])))))/((Y_t[v - 0002]) \\
& / (0.30)) + ((Y_t[v - 0002]) / (0.22)) + (0059) + (Y_t[v - 0001]) + ((Y_t[v - 0006]) - ((0.17) + ((0.63)) / (0.22))) + (Y_t[v - 0001]) + (\sin((Y_t[v - 0002]) / (0.22))) + (\ln1((\cos((wzl((Y_t[v - 0002])))))) - (\cos((Y_t[v - 0002])))) + ((Y_t[v - 0009]) / ((v) / (0.46) * (Y_t[v - 0006])) * (v) + ((Y_t[v - 0010]) * (0059) - (\cos((Y_t[v - 0002])))) * (Y_t[v - 0001]) + ((Y_t[v - 0010]) * (\sin((Y_t[v - 0006])))) + (\sin(\sin(0094) + ((0.25)))) + (\cos((Y_t[v - 0002])))) + (\sin(Y_t[v - 0001])))
\end{aligned} \tag{2.23}$$

mit v als Laufvariable¹⁰. Die beiden Terme wzl und $\ln1$ sind geschützte Rechenoperationen. Der Wurzelterm wzl und der natürliche Logarithmus $\ln1$ sind dahingehend geändert worden, daß nur Absolutwerte als Argumente in diese Rechenoperationen eingehen. Darüber hinaus wurde $\ln(0) \equiv 0$ gesetzt¹¹. Der Vorteil gegenüber der LISP Schreibweise ist, daß der Anwender die - wenn auch sehr komplexe - Gleichung wesentlich leichter nachvollziehen kann. Im Anhang B finden sich die Resultate der übrigen Prozesse inklusive der zugrunde liegenden Schätzgleichungen von GP GAUSS. Bei den bisher angewendeten dynamischen Prozessen handelte es sich um deterministische Systeme. In einem weiteren Artikel schätzt Kaboudan [72] unter anderem dynamische Systeme, die zusätzlich eine stochastische Komponente berücksichtigen. Die dort dargestellten Ergebnisse wurden mit GPQuick generiert und zeigen die Auswertungen des jeweils besten Laufs. Die Güte der genetisch entwickelten Regressionsmodelle wurde dieses Mal anhand der normalisierten durchschnittlichen quadratischen Fehlerabweichung NMSE berechnet:

$$NMSE = MSE / \sigma_Y^2 \tag{2.24}$$

mit σ_Y^2 als die Varianz der abhängigen Werte. Bei den zusätzlichen Gleichungen handelt es sich um zwei deterministische aber komplexere Prozesse von Tong (SH) sowie Mackey & Glass (MG), um zwei nichtlinear-stochastische Modelle von Granger & Andersen (BL) und Liu (TAR), und eine mit UR bezeichnete Zeitreihe, die anhand eines Zufallsgenerators generiert wurde.

Tong [123] (1990) - SH:

$$Y_t = \begin{cases} 2.5Y_{t-1} + x_t & \text{wenn } |Y_{t-1}| \leq 2 \\ 0.85Y_{t-1} & \text{wenn } |Y_{t-1}| > 2 \end{cases}$$

¹⁰ $Y_t[v - 0001]$ entspricht zum Beispiel die um eine Periode verzögerte Zeitreihe von Y_t .

¹¹Siehe auch Kaboudan [72].

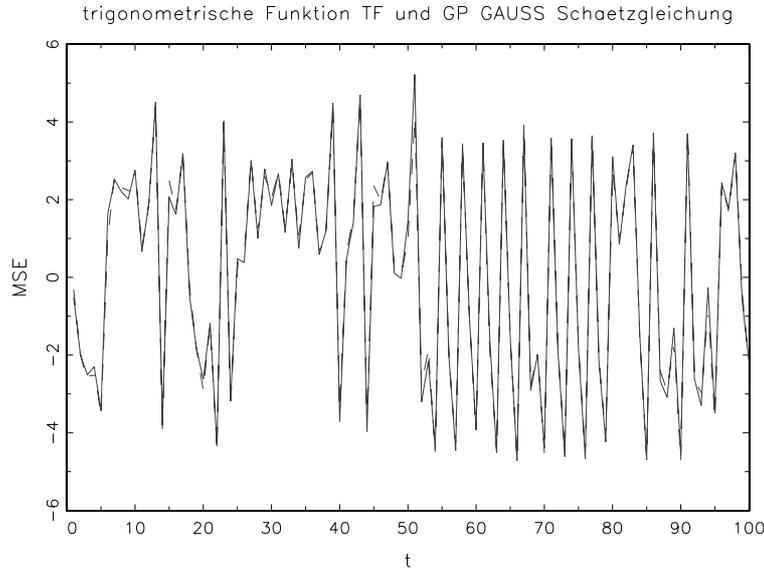


Abbildung 2.7: Darstellung der originalen Zeitreihe der trigonometrischen Funktion TF sowie die geschätzte Zeitreihe durch GP GAUSS.

mit

$$x_t = \begin{cases} -1 & \text{wenn } t = \text{ungerade} \\ 1 & \text{wenn } t = \text{gerade} \end{cases} \quad (2.25)$$

Mackey-Glass [83] Gleichung (1977) - MG:

$$\dot{Y} = \frac{0.2Y_{t-30}}{1 + Y_{t-30}^1} + 0.1Y_t \quad (2.26)$$

Bilineares System von Granger & Andersen [52] aus dem Jahr 1978:

$$Y_t = 0.3Y_{t-1} + 0.6\epsilon_t Y_{t-1} + v_t \quad (2.27)$$

wobei $v_t = N(0, 1)$.

Grenzwert Modell TAR („**t**hreshold **a**utoregressive“) in Liu et al. [79] (1992):

$$Y_t = \begin{cases} 0.9Y_{t-1} + v_t & \text{wenn } |Y_{t-1}| \leq 1 \\ -0.3Y_{t-1} + v_t & \text{wenn } |Y_{t-1}| > 1 \end{cases} \quad (2.28)$$

Für die symbolischen Regressionen wählte Kaboudan 1000 Gleichungen, 100 Generationen und 100 Läufe. Tabelle 2.6 stellt die Ergebnisse von Kaboudan und die von TSGP sowie GP GAUSS gegenüber. Erneut verdeutlichen die Resultate, daß das GP GAUSS

bessere Ergebnisse produziert. Im Vergleich zum GPQuick weist GP GAUSS bei sämtlichen Systemen eine Verbesserung der NMSE Werte um mindestens 50% auf. Nur bei dem mit TAR bezeichneten Grenzwertmodell sind die Resultate mit einer Verbesserung von 7% nur marginal besser. Die Ergebnisse des TSGP sind insgesamt betrachtet recht durchwachsen. Lediglich bei der Mackey-Glass Gleichung hat das TSGP die beste Regressionsgleichung gefunden. Im Vergleich zum GPQuick weist es ansonsten nur bei den weniger komplexen Systemen wie die Ozaki Gleichung oder Logistic Map gute Ergebnisse auf. Die Datengenerierungsprozesse der komplexeren Gleichungen und vor allem solche mit einem stochastischen Term werden mit Ausnahme der Mackey-Glass Gleichung relativ schlecht approximiert. Kaboudan betont in einem seiner Artikel die Problematik, daß das

DGP:	GPQuick	TSGP	GP GAUSS
noise-free models			
OZ:	$1E - 08$	$4E - 9$	$5E - 09$
LG:	0.01	0.007	0.002
HN:	0.01	0.028	0.003
SH:	0.10	0.602	0.021
MG:	0.48	0.112	0.242
non-linear stochastic models			
BL:	0.29	0.748	0.145
TAR:	0.48	0.755	0.442
pseudo-random data			
UR:	0.67	3.719	0.132

Tabelle 2.3: Die Ergebnisse des GPQuick wurden aus dem Artikel von Kaboudan [72] entnommen, während die Resultate von TSGP mit Hilfe der frei verfügbaren Software berechnet wurden.

Genetische Programm eine große Anzahl von Läufen benötigt, um einige wenige qualitativ gute Modelle zu finden. Die Untersuchungen mit GP GAUSS bestätigen diese Beobachtung. Mit zunehmender Komplexität der dynamischen Systeme nimmt die Anzahl guter Regressionen tendenziell ab, wie man den Histogrammen im Anhang B entnehmen kann. Es bleibt festzuhalten, daß das Genetische Programm in der Lage ist, sehr gute In-Sample Schätzungen zu erzielen. Das GP GAUSS liefert insgesamt gesehen die besten Ergebnisse, obgleich die Simulationen ungleich zeitintensiver sind¹².

¹²Weitere statistische Auswertungen sowie graphische Darstellungen und die mathematischen Funktionen der besten Läufe der beiden Tabellen sind im Anhang B aufgelistet.

2.7 Sensitivitätsanalyse

Bevor in den nächsten Kapiteln das Genetische Programm auf verschiedene wirtschaftspolitische Modelle angewendet wird, soll zuvor anhand einer Sensitivitätsanalyse herausgefunden werden, von welchen Faktoren in welchem Ausmaß die Ergebnisse des Genetischen Programms beeinflusst werden. Die Literaturrecherche hat ergeben, daß auf diesem Gebiet bisher relativ wenige Untersuchungen durchgeführt wurden. Das Buch von Banzhaf et al. [5] gibt einige Empfehlungen für die Einstellungen der Anfangsparameter ohne genaue Untersuchungsergebnisse zu präsentieren. Lediglich der Aufsatz der Autoren Chen, Kuo und Shieh [21] (2002) stellt eine umfassendere Analyse dar. Neben diesen Literaturquellen gibt es vereinzelte Erkenntnisse, die im Zusammenhang mit problemorientierten Anwendungen gewonnen wurden. Die Ergebnisse von Chen et al. basieren auf der bereits vorgestellten logistischen Map. Die folgende Sensitivitätsanalyse wird anhand des komplexeren mit SH bezeichneten Prozesses (Tong, 1990) durchgeführt. Ausgehend von den in der Tabelle 2.1 aufgeführten Werten werden *ceteris paribus* einzelne Parameter variiert und ihr Einfluß auf die Fitneßentwicklung überprüft. Die in dem Abschnitt vorgestellte Funktions- und Terminalmenge wird ebenfalls als Ausgangspunkt der Untersuchungen übernommen. Die Zeitreihe SH steht als Variable in der Terminalmenge mit bis zu zwölf Zeitverzögerungen für das Genetische Programm zur Verfügung. Die Sensitivitätsanalyse umfaßt die folgenden Bereiche:

- Variation der Funktionsmenge
- Unterschiedliche Gewichtung genetischer Operationen
- Variation der Populationsgröße und der Anzahl der Generationen
- Variation der Länge der Gleichungen in der Anfangspopulation
- Variation der maximal erlaubten Länge der Gleichungen
- Selektionsmethode für genetische Operationen
- Verwenden verschiedener Fitneßkriterien
- Ansatz zur Analyse der Building Blocks
- Implementieren eines mit Backpropagation bezeichneten Optimierungsverfahrens

Jede Kurve oder Balken der folgenden Graphiken repräsentiert die durchschnittliche Fitneßentwicklung von 100 Läufen mit jeweils 200 Generationen und einer Populationsgröße von 100 Gleichungen, so daß insgesamt pro Kurvenverlauf 2 Millionen Gleichungen analysiert wurden. Falls von dieser Konstellation abgewichen wird, werden die entsprechenden Parametereinstellungen erwähnt. Aufgrund der großen Datenmengen haben die nachstehenden Resultate eine gute Aussagekraft.

2.7.1 Variation der Funktionsmenge

Zu Beginn der Sensitivitätsanalyse soll anhand unterschiedlicher Funktionsmengen überprüft werden, ob eine größere Diversifikation an Funktionsoperatoren oder bestimmte Kombinationen zu Verbesserungen der Fitneßentwicklung führen. Hierzu wurden insgesamt zehn verschiedene Funktionsmengen betrachtet, die in folgende drei große Gruppen eingeteilt werden können: Variation der elementaren Funktionsoperatoren ohne und mit trigonometrische Funktionen sowie unterschiedliche Kombinationen der trigonometrischen Funktionsoperatoren unter Berücksichtigung der elementaren Funktionsoperatoren. Die nachstehende Tabelle 2.4 führt die einzelnen Funktionsmengen im Detail auf. Die beiden

Gruppe:	
F_1	= { + }
F_2	= { + - }
F_3	= { * / }
F_4	= { + - * / }
F_5	= { + abs \sqrt{n} sin cos ln exp }
F_6	= { + - abs \sqrt{n} sin cos ln exp }
F_7	= { * / abs \sqrt{n} sin cos ln exp }
F_8	= { + - * / abs \sqrt{n} sin cos ln exp }
F_9	= { + - * / abs sin cos }
F_{10}	= { + - * / abs \sqrt{n} ln exp }

Tabelle 2.4: Untersuchte Funktionsmengen

Graphiken 2.8 und 2.9 bilden die Fitneßwerte der Funktionsmengen einmal als Balken- und zum anderen als Kurvendiagramm ab. Beim Balkendiagramm sind die durchschnittlichen Fitneßwerte der einhundert Läufe abgetragen. Während die erste Gruppe eindeutig signalisiert, daß eine Ausweitung der Funktionsoperatoren eine Verbesserung der Fitneßentwicklung hervorruft, widersprechen die Ergebnisse der zweiten Gruppe dieser Vermutung. Insgesamt weist die Gruppe F_4 , die den dynamischen Prozeß ausschließlich mit den einfachen Plus-, Minus-, Multiplikations- und Divisionsoperatoren geschätzt hat, das absolute Minimum auf. Unter Hinzuziehung trigonometrischer und anderer Funktionen verschlechtert sich die durchschnittliche Fehlerabweichung wieder. Eine möglichst große Auswahl an Funktionsoperatoren geht daher nicht automatisch mit einer Verbesserung der Ergebnisse einher. Betrachtet man die Funktionsoperatoren des zu approximierenden SH-Prozesses, so fällt auf, daß dieser lediglich Plus-, Minus- und Multiplikationsoperatoren verwendet und am besten mit der Funktionsmenge F_4 übereinstimmt. Die Funktionsmengen F_8 bis F_{10} enthalten ebenfalls diese Funktionsoperatoren und zeigen die drei weiteren besten In-Sample Schätzungen auf. Ein Grund könnte daher sein, daß eine zu differenzierte Funktionsmenge irrelevante Variablen enthält und somit das Auffinden optimaler

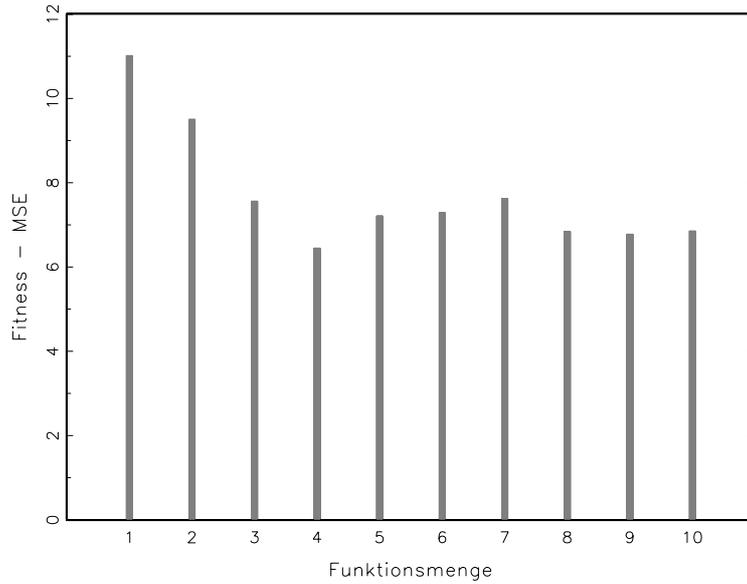


Abbildung 2.8: Funktionsmengen 1 bis 10

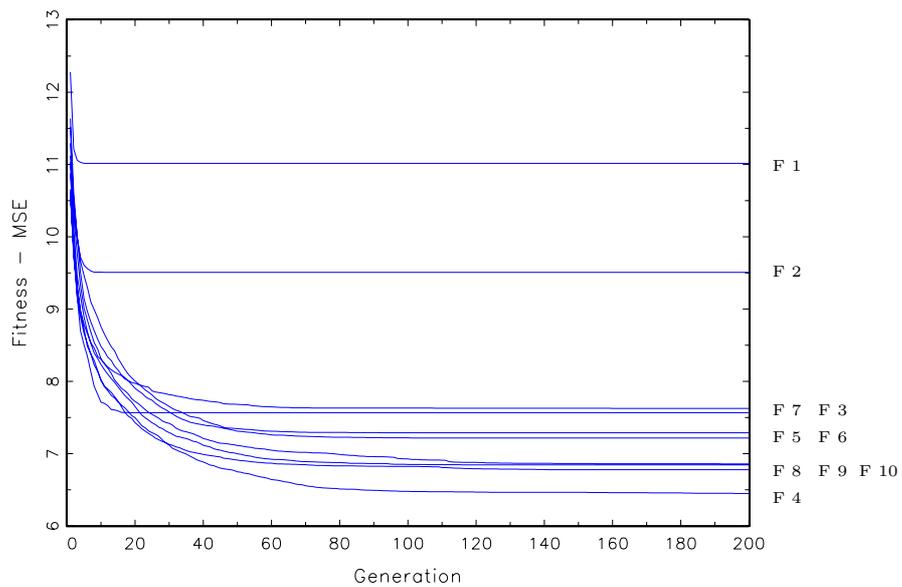


Abbildung 2.9: Funktionsmengen 1 bis 10

Lösungen erschwert. Diese Vermutung behaupten auch Banzhaf¹³ et al. sowie Chen¹⁴ et al.. Während Banzhaf et al. zu der Aussage gelangen, daß kleine aber problemorientierte Funktions- und Terminalmengen vorzuziehen sind, ohne konkrete Ergebnisse darzulegen, bestätigten Chen et al. diese Vermutung anhand einer Sensitivitätsanalyse für die

¹³Vgl. Banzhaf et al. [5], S. 336.

¹⁴Vgl. Chen et al. [21], S. 68 ff.

Terminalmenge. Unter bewußter Einbeziehung irrelevanter Parameter zeigen sie, daß die Ergebnisse erheblich schlechter ausfallen, als wenn die Terminalmenge problemorientiert ausgewählt wird und bereits komplexere Funktionsterme des Datengenerierungsprozesses besitzt. In einem weiteren Experiment zeigen sie, daß sich unter Verwendung zusätzlicher irrelevanter Variablen die Ergebnisse stetig verschlechtern. Der Vorteil einer Berücksichtigung komplexerer Funktionsterme, die den Datengenerierungsprozeß bestimmen, belegen auch die Untersuchungen von Chidambaran et al. [24] (2002). Ihre genetisch entwickelten Option-Preis-Modelle schneiden in ihren Untersuchungen besser ab als die Black & Scholes Formel. Dieses Ergebnis erzielen sie mit einer Terminalmenge, die die Black & Scholes Formel als Parametervariable aufnimmt.

Die Abbildung 2.9 stellt die Kurvenverläufe der einzelnen Simulationsläufe dar und zeigt den typischen hyperbolischen Charakter der Fitneßentwicklung genetisch entwickelter Modelle. Dabei erfolgt die Konvergenz der endgültigen Fitneßgüte in den ersten 50 Generationen.

2.7.2 Unterschiedliche Gewichtung genetischer Operationen

Die folgenden Untersuchungen sollen die Bedeutung der Crossover- und Mutationsoperationen durch unterschiedliche Gewichtungen herausfinden. Bei den Mutationsoperationen handelt es sich um den weiter oben beschriebenen Typ A, dem Substituieren von Operanden durch Operanden bzw. Operatoren durch Operatoren. Es wurden insgesamt elf Modelle simuliert, deren Crossoveranteile schrittweise um jeweils zehn Prozentpunkte beginnend bei null Prozent erhöht wurden. Die Anteile der Crossover- und Mutationsoperationen addieren sich jeweils auf einhundert Prozent. Vergleicht man in Abbildung 2.11 die Modelle mit null und einhundert Prozent Crossoveranteil miteinander, so erkennt man, daß Modelle, die ausschließlich Mutationsoperationen berücksichtigen, wesentlich schlechter abschneiden als Modelle, die allein Crossoveroperationen verwenden. Daraus kann man schließen, daß Crossoveroperationen in einem erheblich stärkeren Ausmaß zum Auffinden optimaler Lösungen beitragen als die Mutationsoperatoren. Die Mutationen sind aber nicht bedeutungslos. Das beweisen die Modelle, die beide genetischen Operationen verwenden. Wie man der Abbildung entnehmen kann, nimmt die Fehlerabweichung bei der Konstellation von 60% Crossover- und 40% Mutationsoperationen ihren minimalen Wert an. Insgesamt betrachtet, ist bereits ein kleiner Anteil an Crossoveroperationen ausreichend, um die Fitneß erheblich zu verbessern. Zu diesem Fazit kommen auch Banzhaf¹⁵ et al. und geben als Daumenregel 90% Crossover und 10% Mutationen an. Die Abbildung kann diesen Vorschlag zwar nicht ganz bestätigen, allerdings liegen die Ergebnisse bei einem Crossoveranteil zwischen 60% und 100% sehr dicht beieinander. Das Ergebnis ist nicht überraschend, da Mutationen lediglich kleine Veränderungen in den Gleichungen verursachen, indem sie einen Knotenpunkt durch einen anderen ersetzen. Crossoveroperationen sind hingegen sehr viel flexibler. Kleine oder große Funktionsbereiche können durch kleine oder große Funktionsbereiche ausgetauscht werden. Während die genetische Operation des Crossovers große Veränderungen bewirkt, sorgt die Mutation vielmehr für die Feinabstim-

¹⁵Vgl. Banzhaf et al. [5], S. 337.

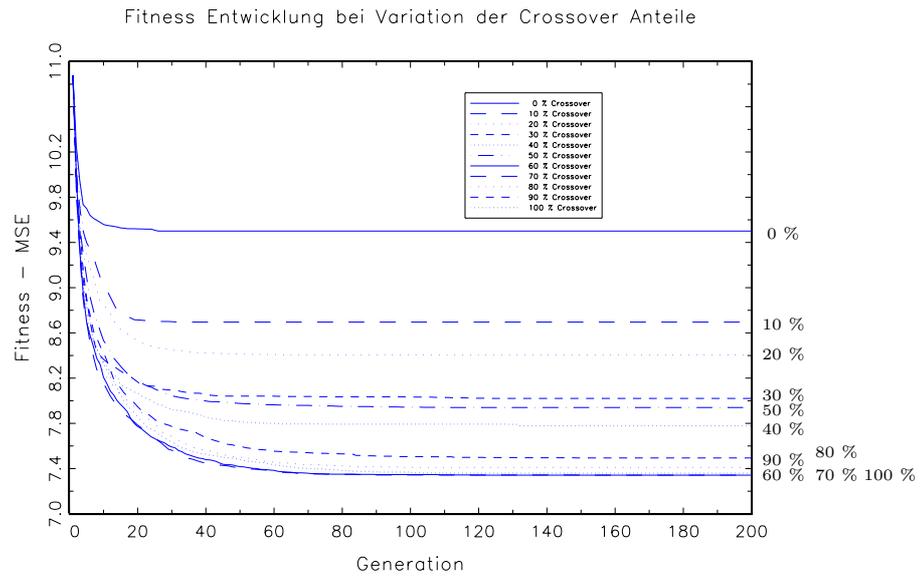


Abbildung 2.10: Variation des Crossoveranteils zwischen 0 und 100%

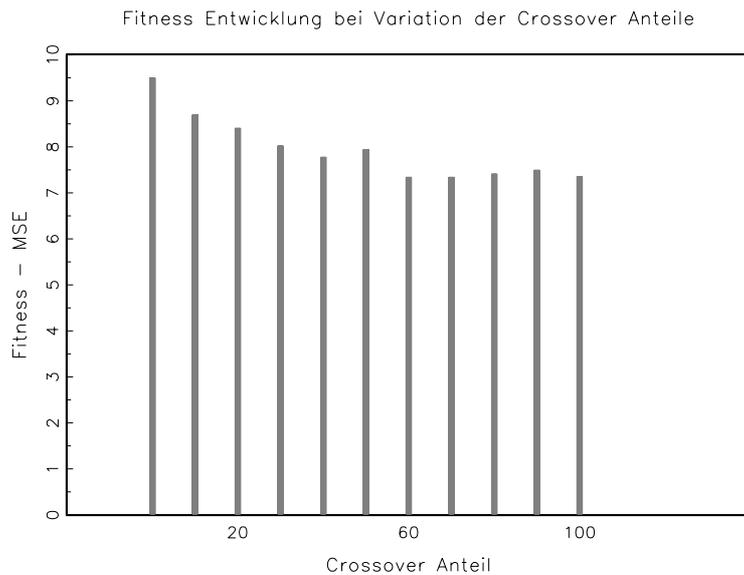


Abbildung 2.11: Variation des Crossoveranteils zwischen 0 und 100%

mung. Dieses Ergebnis wird im Abschnitt der Building Blocks untermauert. Dort wird die Größe des ausgetauschten Teilbereichs in Bezug zu seiner verursachten Fitneßveränderung dargelegt. In ihren Untersuchungen kommen Chen et al. [21] zu den gleichen Aussagen. Zusätzlich zeigen sie, daß Mutationstyp B für komplexere Systeme besonders wichtig ist. Dagegen ist der Reproduktionsoperator, der die besten Lösungen dupliziert, als zusätzliche genetische Operation kaum von Bedeutung. Da sich dieses Resultat ebenfalls während des Programmierens von GP GAUSS zeigte, wurde diese Option eliminiert.

2.7.3 Variation der Populationsgröße und der Anzahl der Generationen

Mit der Variation der Populationsgröße und der Anzahl der Generationen erfolgt eine Variation der Suchintensität. Größere Populationen bieten den Vorteil eines größeren genetischen Pools aus denen eine größere Vielfalt an neuen Lösungen kreiert werden können. Der Nachteil stellt die überproportional ansteigende Rechnerzeit dar, so daß eine möglichst optimale Kombination dieser drei Faktoren mit der Rechnerzeit als Budgetierungsrestriktion erreicht werden soll:

$$\text{Rechnerzeit} = (\text{Anzahl der Generationen}) * (\text{Populationsgröße})$$

Die Ergebnisse in der Literatur fallen recht unterschiedlich aus. Die Autoren Gathercole & Ross [47] (1997) bevorzugen kleine Populationen über viele Generationen im Vergleich zu großen Populationen über wenige Generationen. In ihren Untersuchungen variierten sie die Populationsgrößen zwischen $Pop = 50$ und $Pop = 10000$ und erzielten das beste Ergebnis mit einer Populationsgröße von $Pop = 200$. Aufgrund der benötigten Rechenzeit bevorzugten sie aber einen kleineren Umfang von $Pop = 50$ Gleichungen, was sie wahrscheinlich wegen der mittlerweile erheblich verbesserten Rechengeschwindigkeit zugunsten der Populationsgröße von $Pop = 200$ revidieren würden. Banzhaf et al. behaupten dagegen, eine annähernd proportionale Verbesserung der Fitneß bei Zunahme der Populationsgröße vorzufinden¹⁶. Insgesamt favorisieren sie eher größere Populationen und beziffern einen Umfang von $Pop = 1000$ Gleichungen als einen guten Startwert. Bei Zeitreihen mit einer stärker ausgeprägten Stochastik bevorzugen die Autoren allerdings kleinere Populationsgrößen. Koza¹⁷ wiederum nennt eine Größe von $Pop = 500$ als angebracht. Die Untersuchungen von Chen et al. können keine proportionale Verbesserung bei einer linearen Erhöhung der Populationsgrößen feststellen¹⁸. Vielmehr finden sie eine abnehmende Grenzrate der Verbesserung. Um die Sensitivität hinsichtlich der Populationsgröße zu

Population	Anzahl der Läufe	Anzahl Generationen	MSE- ϕ	Varianz- ϕ	MSE-beste	Rechenzeit in Stunden
100	100	1000	6.84	3.05	2.65	26.84
500	20	1000	4.25	2.22	2.02	24.33
1000	10	1000	3.59	0.43	2.38	25.46

Tabelle 2.5: Pro Lauf wurden stets 1000 Generationen erzeugt; MSE- ϕ = durchschnittliches MSE, Varianz- ϕ = durchschnittliche Varianz, MSE-best = MSE der besten Funktion.

analysieren, simulierte GP GAUSS drei Programme mit Populationsgrößen von 100, 500 und 1000 sowie 1000 Generationen. Die Anzahl der Läufe wurde so angepaßt, daß GP GAUSS insgesamt jeweils 10 Millionen Gleichungen generierte. Die in Tabelle 2.5 dargestellten Ergebnisse zeigen, daß eine Populationsgröße von $Pop = 500$ die besten Resultate liefert. Zwar weist das Programm mit $Pop = 1000$ den besten Durchschnittswert auf, aber

¹⁶Vgl. Banzhaf [5], S.335.

¹⁷Vgl. Koza [75].

¹⁸Vgl. Chen et al. [21], S. 62 ff.

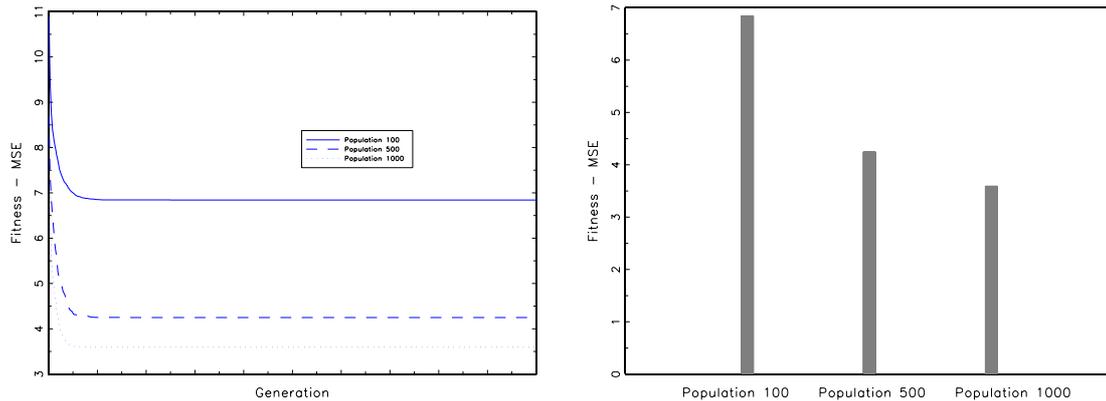


Abbildung 2.12: Variation der Populationsgröße

$Pop = 500$ erzeugte die beste Regressionsgleichung bei minimaler Rechenzeit. Die Ergebnisse unterstützen somit den Vorschlag von Koza. Darüber hinaus stimmen die Resultate mit denen von Chen et al. überein, da die durchschnittlichen MSE Werte sich nicht linear verbessern. Auch ein Blick auf die Graphik 2.12 zeigt, daß die Abstände der Kurven zwischen $Pop = 100$ und $Pop = 500$ im Vergleich zum Abstand $Pop = 500$ und $Pop = 1000$ nicht proportional ausfallen. Weiterhin zeigen die Fitneßkurven, daß eine Ausweitung der Anzahl auf 1000 Generationen wegen der Konvergenz nur wenig sinnvoll erscheint.

2.7.4 Variation der Länge der Gleichungen in der Anfangspopulation

In diesem Abschnitt soll überprüft werden, welchen Einfluß die Anfangspopulation auf das Endergebnis ausüben kann. Dazu wird die Länge der Gleichungen, die in der Anfangspopulation erzeugt wird, variiert. Anfangspopulationen mit sehr langen Gleichungen bieten den Vorteil, daß eine große Menge an genetischem Material für die Folge-Generationen zur Verfügung steht. Andererseits enthalten sie viele unnötige Terme, die das Ergebnis zunächst verschlechtern und folglich abgestoßen werden müssen. Bei kurzen Anfangsgleichungen verhält es sich genau umgekehrt. Aus den beiden Funktionsmengen und der Terminalmenge wählt das Genetische Programm mit Hilfe eines Zufallsgenerators die einzelnen Elemente und baut diese zusammen. Aufgrund des Programmaufbaus ist es nicht möglich, eine vorher festgelegte Knotenanzahl exakt zu erreichen. Es können aber Bereiche angegeben werden, die die minimale und maximale Länge einer Gleichung bestimmen. Legt man die Länge auf K Knoten fest, so wird die tatsächliche Länge in dem Bereich $2 * K - 1 \leq Knoten \leq 2 * (K - 1) + K$ liegen. Tabelle 2.6 offenbart, daß sich die durchschnittlichen Fitneßwerte in der ersten Generation mit zunehmender Länge der Anfangsgleichungen verschlechtern. Das Ergebnis für $K = 30$ mit einer durchschnittlichen Knotenlänge von 73.5 in der Anfangspopulation weist mit einem MSE-Wert von 12.06 größere Fehlerabweichungen als für $K = 2$ auf. Auch die Ergebnisse nach 200 Generationen hängen negativ von der Knotenlänge der Anfangsgleichungen ab. Auf der nächsten Seite ist sowohl die Fitneß- als auch die Knotenentwicklung abgebildet. In Bezug auf die Fitneßentwicklung in Abbildung 2.14 konvergieren alle Modelle ähnlich schnell zu ihrem Endwert

K	K-Minimum	K-Maximum	ϕ -Knoten	ϕ -MSE	
				Gen = 1	Gen = 200
2	3	4	3.5	10.69	5.31
5	9	13	11.0	11.30	6.12
10	19	28	23.5	11.88	7.44
20	39	58	48.5	11.97	8.24
30	59	88	73.5	12.06	8.65

Tabelle 2.6: ϕ -Knoten = durchschnittliche Knotenanzahl in Generation eins; ϕ -MSE = durchschnittlicher Fitneßwert.

und weisen den typischen Verlauf auf. Die oben aufgestellte Vermutung wird in Abbildung 2.13 bestätigt. Die Knotenanzahl großer Anfangsgleichungen reduziert sich zunächst in den ersten Generationen, um sich dann der maximal erlaubten Länge zu nähern.

2.7.5 Variation der maximal erlaubten Länge der Gleichungen

In dem Grundmodell wurde die maximale Länge der Schätzgleichungen auf 100 Knoten festgelegt. Wie verändern sich die Ergebnisse bei Verschieben dieser Grenze sowohl nach oben als auch nach unten? Die beiden Graphiken in Abbildung 2.15 belegen eine minimal stetige Verbesserung bei zunehmender Knotenanzahl. Bei Zulassen größerer Gleichungen verstärkt sich der bereits weiter oben angesprochene 'Bloating'-Effekt. Das Ansteigen redundanter Funktionsbereiche bei zunehmender Länge mathematischer Gleichungen ist zunächst negativ anzusehen, da aufgrund der Komplexität der gefundenen Gleichungen, ein Herausfiltern der redundanten Bereiche sehr aufwendig ist. Der Vorteil der längeren Gleichungen liegt in der größeren Vielfalt an austauschbaren genetischen Teilbereichen im Zusammenhang vor allem mit den Crossoveroperationen. Nützliche Teilbereiche werden in der Literatur mit 'Building Blocks' umschrieben. Ein weiteres Problem größerer Gleichungen besteht im sogenannten Overfitting. Bei symbolischen Regressionen ist dieses Problem unbedeutend. Bei der Entwicklung von Prognosemodellen wie beispielsweise für Finanzmarktzeitreihen stellt es allerdings ein Problem dar, für das es bis jetzt keine Lösung gibt. Beim Overfitting wird die vorliegende Zeitreihe zwar immer besser geschätzt, wendet man aber die genetisch entwickelten Modelle für Prognosen an, muß man feststellen, daß sie zusehends schlechter werden. Das Problem liegt darin, daß das Programm beginnt zu versuchen, die in den Daten vorhandene Stochastik erklären zu wollen. Bei zukünftigen Anwendungen wird die Länge auf 100 Knoten beschränkt, auch wenn das Problem des Overfitting damit nicht beseitigt ist.

2.7.6 Selektionsmethode für genetische Operationen

Für die genetischen Operationen Crossover und Mutation müssen zwei bzw. eine Elternfunktion zufällig ausgewählt werden. Im Grundmodell ist diese Wahrscheinlichkeit abhängig vom Fitneßrang einer Gleichung. Sollen beispielsweise anhand der besten n Glei-

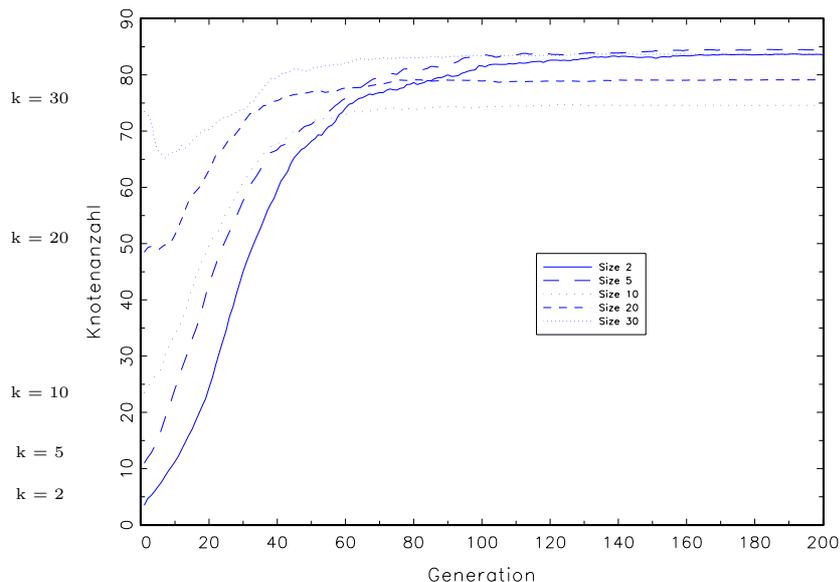


Abbildung 2.13: Variation der Anfangspopulation - Knotenentwicklung

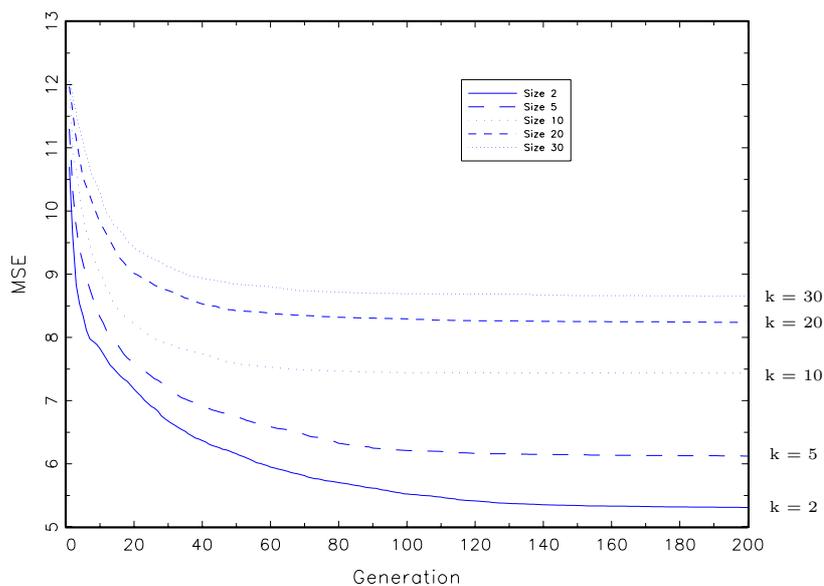


Abbildung 2.14: Variation der Anfangspopulation - Fitneßentwicklung

chungen sämtliche Crossover- und Mutationsoperationen durchgeführt werden, berechnet das Programm zunächst folgenden kumulativen Zahlenvektor:

$$\text{Zahlenvektor } z_i = \sum_{l=1}^i \frac{1}{1+l} \quad 1 \leq i \leq n \quad (2.29)$$

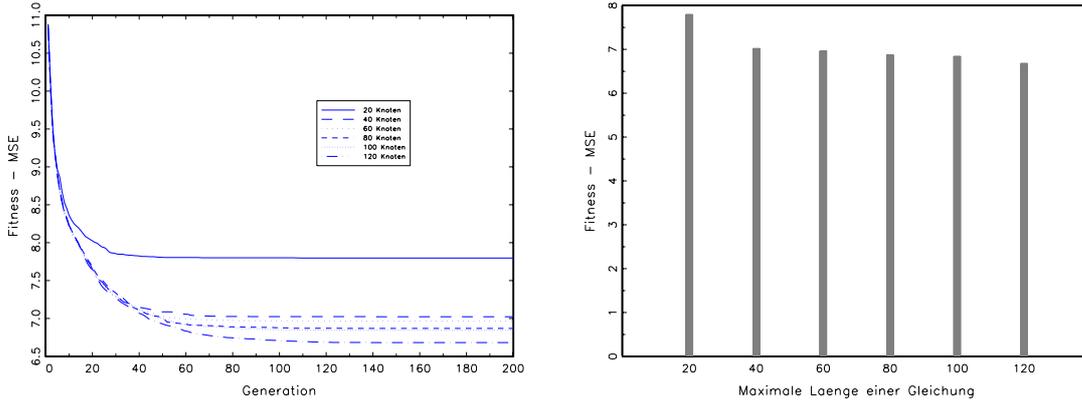


Abbildung 2.15: Variation der Länge mathematischer Gleichungen

Man erhält einen $n \times 1$ Vektor, dessen Einträge Intervallgrenzen ausdrücken. Für $n = 5$ erhält man beispielsweise:

$$\text{Zahlenvektor}(5) = \begin{pmatrix} 0.34 \\ 0.57 \\ 0.75 \\ 0.89 \\ 1 \end{pmatrix} \quad (2.30)$$

Ein Zufallsgenerator zieht eine Zahl zwischen null und eins und vergleicht diese mit den Intervallgrenzen. Der resultierende Intervallindex gibt diejenige Gleichung an, die aus der Menge der (gemäß den Fitneßwerten geordneten) Gleichungen gewählt wird. Liegt zum Beispiel eine Zufallszahl im zweiten Intervall, so wird die zweitfitteste Gleichung aus der Menge der fünf fittesten Gleichungen für eine genetische Operation gewählt. Die Wahrscheinlichkeit einer Gleichung ausgewählt zu werden, nimmt mit Verschlechterung des Rangs ab. Daher bezeichnet man dieses Verfahren in der Literatur auch als proportionale Selektionsmethode. Der Vektor mit den Intervallgrenzen bleibt bei dieser Methode im Zeitablauf aber konstant. Es wurden Modelle untersucht, die die besten 25, 50, 75 oder 100 Gleichungen mit entsprechendem Intervallvektor für genetische Operationen benutzt haben. Die Ergebnisse werden zunächst mit einer leicht modifizierten proportionalen Selektionsmethode verglichen. In Gleichung 2.31 wird der Fitneßrang durch den Fitneßwert ersetzt, so daß sich der kumulative Zahlenvektor folgendermaßen berechnet:

$$\text{Zahlenvektor} \quad z_i = \sum_{l=1}^i \frac{\frac{1}{1+MSE_l}}{\sum_{j=1}^n \frac{1}{1+MSE_j}} \quad 1 \leq i \leq n \quad (2.31)$$

Da die modifizierte Selektionsmethode vom Fitneßwert abhängt, verändert sich der Intervallvektor im Zeitablauf gemäß der Fitneßentwicklung. Bessere Schätzgleichungen mit einer kleineren mittleren quadratischen Fehlerabweichung bekommen eine höhere Auswahlwahrscheinlichkeit zugewiesen. In Abbildung 2.16 sind die Ergebnisse der beiden Methoden gegenübergestellt. Es wird deutlich, daß der Unterschied zwischen einer rang- und einer fitneßabhängigen Selektionsmethode nur geringfügig ausfällt mit leichten Vorteilen einer

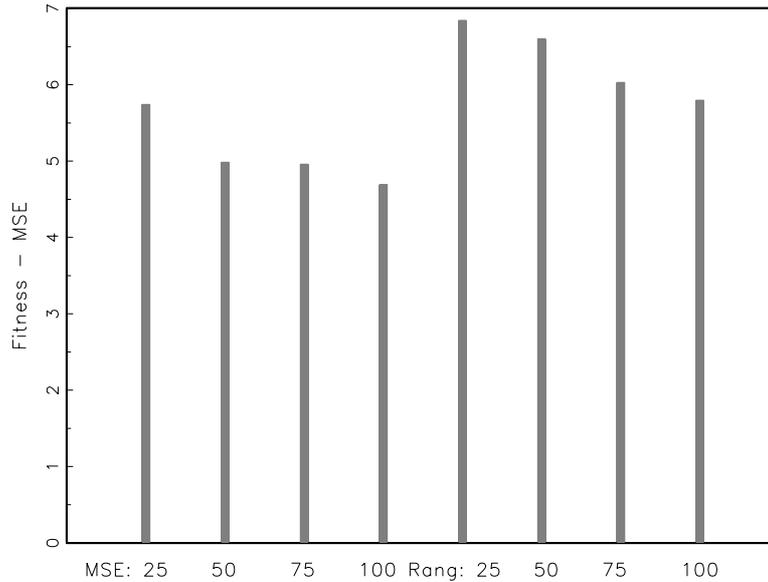


Abbildung 2.16: Vergleich von MSE und Fitneßrang

fitneßbasierten Auswahl. Darüber hinaus erfolgt eine stetige Verbesserung der Ergebnisse bei einer Ausweitung der Menge an Gleichungen unabhängig von der Selektionsmethode.

Chen et al. haben die proportionale Selektionsmethode mit der Tournament Selektionsmethode und einem uniformen, d.h. gleichwahrscheinlichen, Verfahren verglichen. Letzteres folgt nicht mehr dem “survival-of-the-fittest” Prinzip. Beim Tournament Verfahren¹⁹ werden aus der Population zufällig zwei oder mehrere Gleichungen gezogen. Die Gleichung mit der besten Fitneß gewinnt das Turnier und wird anschließend in einer Tournament Gruppe platziert. Der Auswahlprozeß wiederholt sich solange, bis die Tournament Gruppe eine bestimmte Größe erreicht hat. Anhand der Tournament Gruppe werden wiederum die Gleichungen für genetische Operationen mit einer gleichverteilten Auswahlwahrscheinlichkeit zufällig gezogen. Der Vorteil dieser Selektionsmethode besteht darin, daß die Menge an Gleichungen, die für genetische Operationen herangezogen werden, in jeder Generation sehr unterschiedlich zusammengesetzt ist. Gleichungen mit einer schlechteren Fitneß besitzen gegenüber dem proportionalen Verfahren eine größere Auswahlwahrscheinlichkeit. Chen et al. kommen zu dem Ergebnis, daß die uniforme Selektionsmethode in manchen Fällen im Vergleich zur proportionalen Selektionsmethode gleichgut oder sogar besser ist. Im Vergleich zur uniformen und proportionalen Methode schneidet die Tournamentmethode wiederum besser ab. Anhand einer Tournament Gruppe mit 25 Gleichungen sollte GP GAUSS diese doch recht überraschenden Ergebnisse überprüfen. Für die uniforme Selektionsmethode wurde die Population erneut nach ihrer Fitneß geordnet und die besten 25 Gleichungen wurden für genetische Operationen ausgewählt. Die Selektion einer Gleichung für genetische Operationen aus dieser Gruppe erfolgt mit einer gleichverteilten Wahrscheinlichkeit. Zum Vergleich: die proportionale Selektionsmethode wählt die besten 25 - nach der Fitneß sortierten - Gleichungen, wobei die Selektionswahrscheinlichkeit ei-

¹⁹Siehe auch Goldberg et al. [50] und Drake et al. [31].

	Proportionale Selektion	Tournament Selektion	uniforme Selektion
Mittelwert	6.84	7.72	6.94
Varianz	3.06	2.81	2.61
beste Regression	2.65	3.02	3.12

Tabelle 2.7: Die Werte geben den Mean Squared Error wieder.

ner Gleichung für genetische Operationen vom Rang abhängt. Die Tabelle 2.7 zeigt den Mittelwert und die Varianz von 100 Läufen sowie den Fitneßwert der besten Regression. Demnach liefert die proportionale Selektionsmethode die besten Resultate. Während der durchschnittliche MSE Wert im Vergleich zur uniformen Selektionsmethode nur leichte und gegenüber dem Tournamentverfahren eindeutige Vorteile aufweist, weicht der MSE Wert der fittesten Gleichung eindeutig zugunsten der proportionalen Methode ab. Das Tournamentverfahren kann nur im Vergleich zum uniformen Selektionsverfahren in Bezug auf die beste Regressionsgleichung punkten. Lediglich die Varianz fällt schlechter aus, so daß zwar einerseits sehr gute aber andererseits auch sehr schlechte Läufe hervorgebracht werden. Insgesamt können die Ergebnisse von Chen et al. zumindest für diese Untersuchung nicht bestätigt werden.

2.7.7 Verwenden verschiedener Fitneßkriterien

Nachdem in dem Abschnitt zuvor die Anzahl der besten Gleichungen sowie die Selektionsmethode für genetische Operationen variiert wurden, konzentriert sich dieser Abschnitt auf das Kriterium, nach dem die Gleichungen bewertet und geordnet werden. In den bisherigen Modellen war dies stets die mittlere quadratische Fehlerabweichung. Es berechnet die Güte einer Gleichung, berücksichtigt aber nicht die Länge einer Gleichung. Kürzere Gleichungen sind aber wegen der bereits weiter oben angesprochenen Gründen vorzuziehen. Ein Kriterium, daß die Länge einer Gleichung bestraft, ist zum Beispiel das Informationskriterium von Akaike AIC:

$$AIC = \ln(MSE) + \frac{2 * m}{n} \quad (2.32)$$

mit m = Anzahl der Regressoren und n = Anzahl der Beobachtungswerte. Je mehr Regressoren die Gleichung aufweist, desto größer wird sie bestraft. In meinen Untersuchungen entspricht m der Knotenanzahl, die mit verschiedenen Gewichtungen n in die Bewertung eingeht. Da die Anzahl der Knoten nicht gleichzusetzen ist mit der Anzahl der Regressoren, entsprechen die Berechnungen nur der Idee nach dem Informationskriterium. Der Unterschied ist für die Qualität der Ergebnisse irrelevant, da es zum einen in der Literatur für das Akaike Kriterium generell keine einheitliche Darstellung gibt. Zum anderen stellt es le-

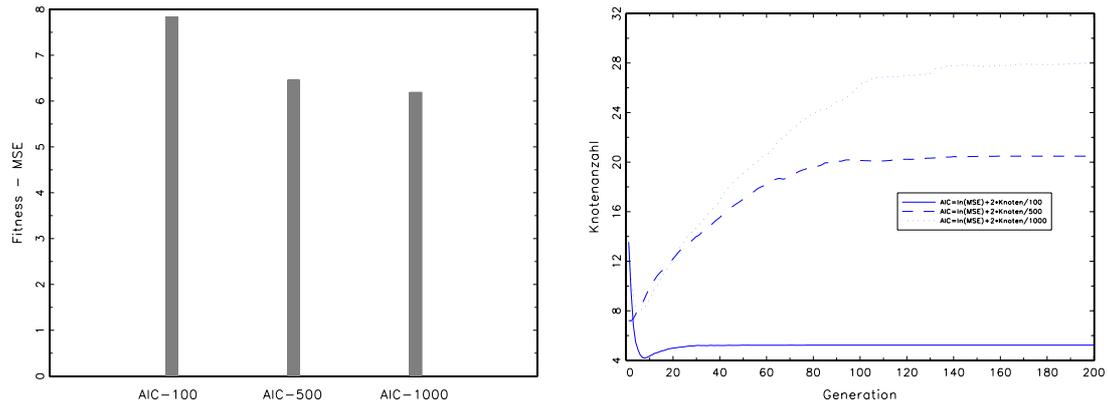


Abbildung 2.17: Verwenden eines alternativen Selektionskriteriums

diglich eine monotone Transformation dar. Wie man der Abbildung 2.17 entnehmen kann, wurden drei Modelle mit Werten für $n = 100, 500$ und 1000 untersucht. Die Abbildung zeigt, daß eine zu starke Gewichtung der Länge ($n = 100$) die Ergebnisentwicklung sehr negativ beeinflusst. Die durchschnittliche Länge der Gleichungen konnte zwar erfolgreich verringert werden, aber zu Ungunsten der Fitneßentwicklung. Die Knotenanzahl bewegt sich nahezu konstant bei 5 Knoten. Geht die Länge einer Gleichung mit einem geringeren Gewicht in die Fitneßbewertung ein, so verbessert sich das Fitneßkriterium kontinuierlich. Je geringer die Gewichtung, desto besser entwickelt sich die Fitneß auf Kosten längerer Funktionen. Zumindest für die In-Sample Analyse ist das MSE dem Informationskriterium vorzuziehen. Da Informationskriterien bei Prognosemodellen eine wichtige Funktion einnehmen, sollen sie in den späteren Kapiteln erneut angewendet werden.

2.7.8 Ansatz zur Analyse der Building Blocks

Als Building Block bezeichnet man den Teilbereich einer Gleichung, der bei genetischen Crossoveroperationen ausgetauscht wird. In der Literatur gibt es nur wenige Studien, die sich mit Building Blocks beschäftigen. Dabei werden sie als Ursache für das erfolgreiche Funktionieren der Crossoveroperationen angegeben. Für die im Folgenden dargestellten Ergebnisse existiert meines Wissens nach noch keine Studie, die die Buildings Blocks auf diese Weise analysiert haben. Die Ergebnisse sollen dazu beitragen, eine Antwort auf die in der Literatur aufgestellten Aussagen zu geben. Neben der Analyse der Crossoveroperationen werden die Auswirkungen der Mutationsoperationen analysiert. Sie stellen sozusagen die minimal möglichen Building Blocks dar.

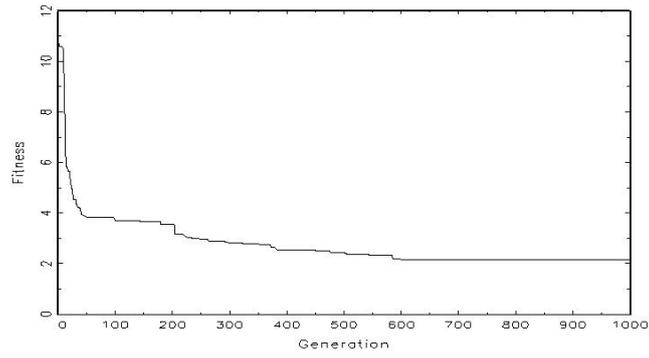
Berechnung der Fitneßänderungen

Die Güte der Building Blocks wird wie folgt untersucht: In jeder Generation sollen die besten 100 Gleichungen für genetische Operationen zur Verfügung stehen. Aus den Gleichungen werden bei Crossoveroperationen die Teilbereiche zufällig ausgeschnitten und die Länge bzw. die Knotenanzahl wird gespeichert. Anhand eines Zufallsgenerators erfolgt

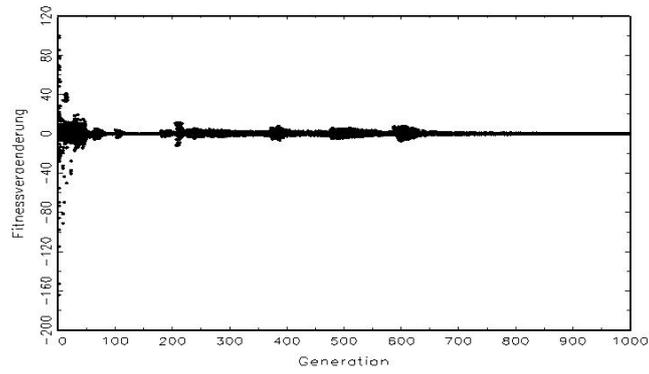
eine proportionale und rangbasierte Auswahl der jeweiligen Gleichungen. Die Indizes werden ebenfalls gespeichert, um den Fitneßwert der ursprünglich ausgewählten Gleichungen sowie die Länge der ausgetauschten Teilbereiche zurückverfolgen zu können. Die Fitneß der neuen Gleichungen wird berechnet und mit den ursprünglichen verglichen. Es werden ausschließlich die Crossover- und Mutationsoperationen in den folgenden Untersuchungen berücksichtigt, die innerhalb der besten einhundert Gleichungen zu einer Fitneßverbesserung geführt haben und somit in der nächsten Generation beibehalten werden. Durch diese Methode kann untersucht werden, ob und in welchem Ausmaß die Buildings Blocks zur Ergebnisverbesserung beigetragen haben. Zudem ist es möglich, die Auswirkungen und die Anzahl erfolgreicher Building Blocks im Zeitablauf darzustellen. Darüber hinaus kann festgestellt werden, ob große Building Blocks tendenziell größere Verbesserungen als kleine mit sich bringen. Letzteres wird in der Literatur häufig behauptet und durch die Tatsache, daß Crossoveroperationen bessere Ergebnisse als Mutationsoperationen liefern (siehe auch Abbildung 2.11), belegt.

Ergebnisse

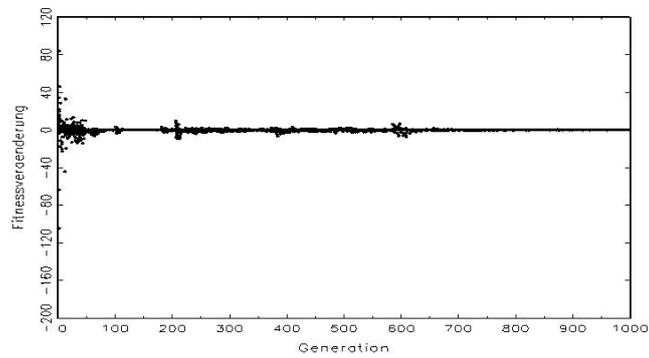
Abbildung 2.18 zeigt die Fitneßentwicklung eines Laufs für 1000 Generationen sowie weitere graphische Analysen. Der Anteil an Crossover- und Mutationsoperationen wurde mit jeweils 50% gleichgewichtet. Aufgrund des sprunghaften Verlaufs zu Beginn der ersten 50 Generationen sowie in den Bereichen um Generation 200, 370 und 600 eignet der Lauf sich sehr gut für die nachfolgenden Untersuchungen. Die im Zeitablauf verursachten Fitneßveränderungen durch Mutationsoperationen und Crossoveroperationen zeigen deutlich, daß bei den oben erwähnten Generationen die Veränderungen größere Ausschläge verzeichnen als in den anderen Generationen. Interessant ist vor allem, daß sowohl Crossover als auch Mutationen ein sehr ähnliches Volatilitätsmuster aufweisen. Aus den Abbildungen lassen sich mehrere Ergebnisse festhalten. Als erstes fällt auf, daß die genetischen Operationen neben positiven auch negative Fitneßveränderungen verursachen. Negative Veränderungen sind möglich, da der neue Fitneßwert trotz Verschlechterung gegenüber der ursprünglichen Gleichung zu den besten 100 Werten gehört und daher in der nächsten Generation beibehalten wird. Die zu Beginn erhaltene Fitneßverbesserung um 100% resultiert aufgrund von gerundeten Zahlen. Gerade in der ersten Generation kann es vorkommen, daß die Gleichungen Fitneßwerte erzeugen, die gegen unendlich tendieren. Wenn eine genetische Operation anschließend eine Gleichung produziert, die einen vernünftigeren Fitneßwert hervorbringt, so ergibt die gerundete Zahl eine Verbesserung von 100%. Des weiteren zeigen die Abbildungen, daß nach jedem Fitneßsprung die Veränderungen anschließend kontinuierlich abnehmen. Um diesen Sachverhalt deutlicher erkennbar zu machen, wurde der Bereich zwischen Generation 450 und 700 vergrößert dargestellt. Betrachtet man Abbildungen 2.19, so erkennt man bei Generation 480 und 580 sprunghafte Fitneßverbesserungen. Im Vergleich dazu zeigen die Fitneßänderungen der Crossover- und Mutationsoperationen einen markanten Anstieg sowohl der positiven als auch negativen Änderungsraten mit anschließender kontinuierlicher Abnahme. Die sprunghafte Verbesserung durch innovative genetische Operationen und der anschließenden Feinabstimmung der gefundenen Gleichung könnten bei der späteren Anwendung von GP GAUSS zur Ge-



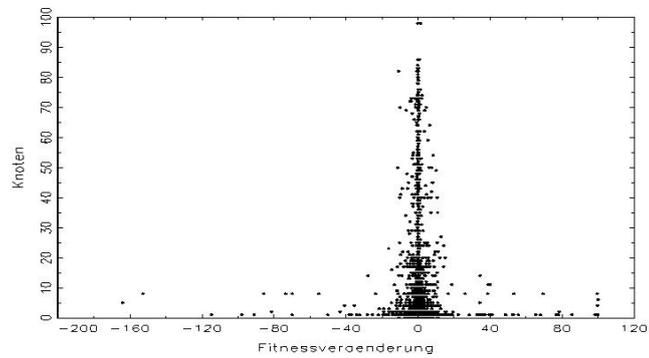
(a) MSE Fitneßentwicklung



(b) Fitneßänderungen: Crossoveroperationen im Zeitablauf

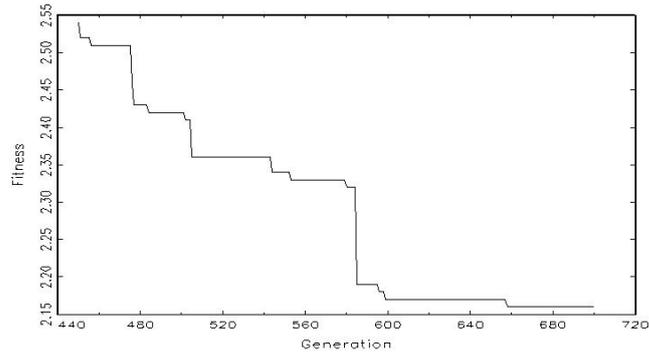


(c) Fitneßänderungen: Mutationsoperationen im Zeitablauf

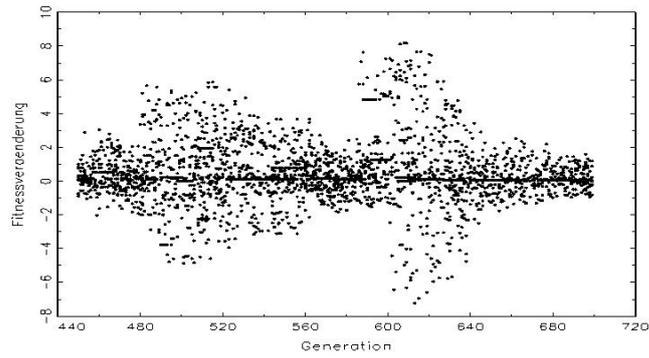


(d) Fitneßänderungen ausgetauschter Teilstränge in %

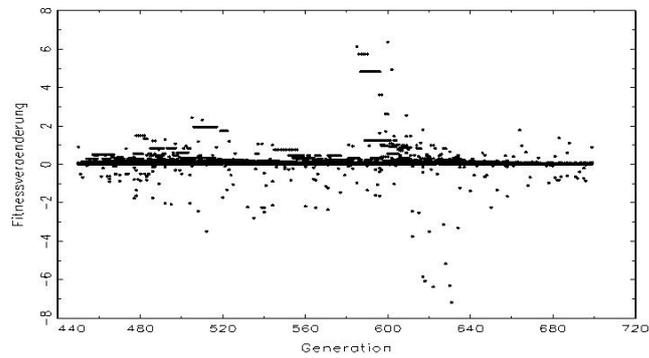
Abbildung 2.18: Fitneßänderungen eines Laufs in %: 1. bis 1000. Generation



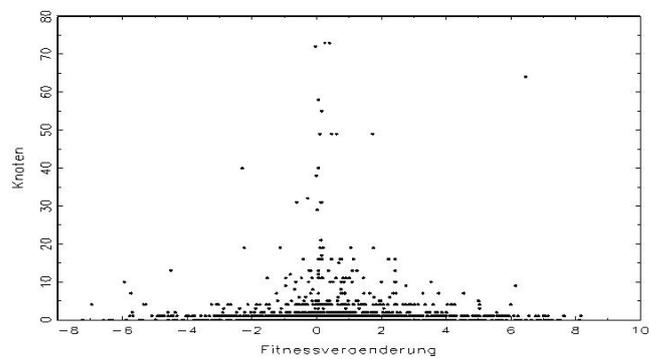
(a) MSE Fitneßentwicklung



(b) Fitneßänderungen: Crossoveroperationen im Zeitablauf



(c) Fitneßänderungen: Mutationsoperationen im Zeitablauf



(d) Fitneßänderungen ausgetauschter Teilstränge in %

Abbildung 2.19: Ausschnitt des obigen Laufs: 450. bis 700. Generation

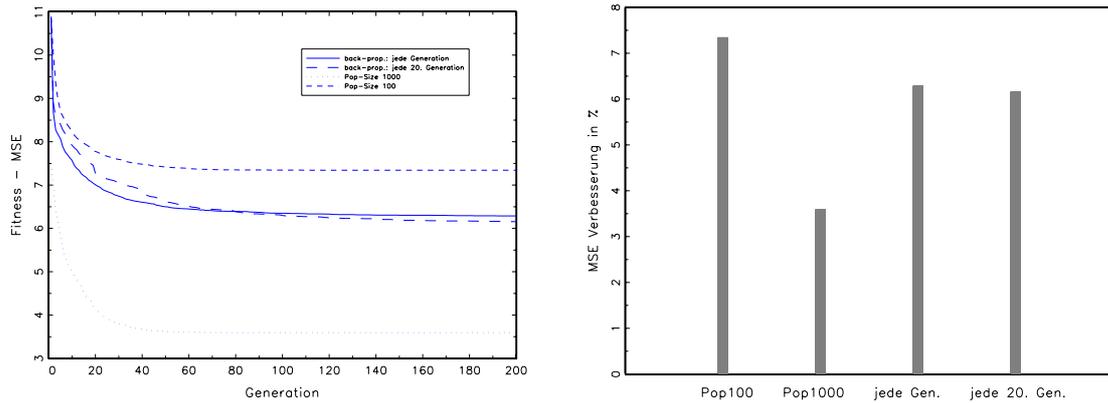
		Lauf 1	Lauf 2	Lauf 3	Lauf 4	Lauf 5	Lauf 6
Crossover	+	7272	6749	3897	5205	5105	4187
	-	4172	4004	2545	3732	2869	2585
Relation		1.74	1.69	1.53	1.39	1.78	1.62
Mutation	+	19873	1802	7075	10653	21017	9868
	-	929	97	527	1044	472	841
Relation		21.39	18.56	13.43	10.20	44.53	11.73
MSE		2.16	2.23	6.40	2.34	6.07	5.14

Tabelle 2.8: Anzahl positiver und negativer Änderungsraten

nerierung von Prognosemodellen auf Finanzmärkten nützlich sein. Vielleicht ist es unter Zuhilfenahme dieser methodischen Analyse möglich, den Zeitpunkt eines eventuellen Overfittings besser abschätzen zu können.

Besonders auffällig ist die nahezu symmetrische Verteilung positiver und negativer Änderungsraten bei beiden genetischen Operationen, wobei die Symmetrie bei der Crossoveroperation stärker ausgeprägt ist. Dies belegen auch die Relationen an positiven und negativen Änderungsraten in Tabelle 2.8. Bei insgesamt 100 000 genetischen Operationen, die jeweils zur Hälfte auf Crossover- und Mutationsoperationen entfallen, weisen Mutationsoperationen 19873 positive und 929 negative Vorzeichen auf. Bei den Crossoveroperationen sind es 7272 respektive 4172. Die Relationen wurden für fünf weitere Läufe berechnet. Die symmetrischen Eigenschaften finden sich in den anderen Läufen ebenfalls wieder. Besonders interessant sind die Relationsgrößen. Bei Crossoveroperationen schwankt die Relation zwischen 1.39 und 1.78. Die Anzahl positiver und negativer Änderungsraten ist wesentlich ausgeglichener als bei Mutationsoperationen, deren Relation zwischen 10.2 und 44.53 schwankt. Betrachtet man die Fitneßwerte der einzelnen Läufe, so kann man eine leichte lineare Abhängigkeit zu der absoluten Anzahl an positiven Änderungsraten bei Crossoveroperationen erahnen. Dies ist bei den Mutationsoperationen nicht der Fall. Der fünfte Lauf besitzt mit 21017 die meisten positiven Änderungsraten bei Vorliegen des zweitschlechtesten Fitneßwertes von 6.07.

Der graphische Vergleich zwischen Crossover und Mutation offenbart weiterhin, daß die prozentualen Verbesserungen der Crossoveroperationen deutlich stärker ausfallen sowohl in den ruhigen als auch bei den volatileren Phasen. Lediglich in den ersten 50 Generationen sind die Änderungsraten ungefähr gleichgroß. Als nächstes wird bei Crossoveroperationen der Zusammenhang zwischen Fitneßveränderung und Knotenlänge eines Teilbereichs untersucht. Die Graphiken 2.18(d) und 2.19(d) bilden dazu die Anzahl der Knotenlänge der ausgetauschten Teilbereiche gegen die Fitneßänderungsrate ab. Auch im Vergleich zu den anderen Graphiken (siehe Anhang) fällt auf, daß im Bereich 1 – 15 Knoten die größten Änderungsraten vorliegen. Da bei Crossoveroperationen Teilbereiche mit einer Länge von einem Knoten ausgetauscht werden können, enthält die Graphik auch Pseudo-Mutationen. Insgesamt zeigen die Graphiken, daß kleinere Building Blocks insgesamt größere Änderungsraten hervorrufen.



(a) MSE Fitneßentwicklung

Abbildung 2.20: Fitneßentwicklung bei Anwenden des Backpropagation Moduls. Die Populationsgröße bei Anwendung des Backpropagation Moduls betrug jeweils 100 Gleichungen.

2.7.9 Implementieren eines mit Backpropagation bezeichneten Optimierungsverfahrens

Die Idee ein zusätzliches Modul in das bestehende Computerprogramm einzubauen, das die Gewichte einer Schätzgleichung optimiert, entsprang aus dem Bereich der künstlichen Neuronalen Netze. Der Backpropagation Algorithmus arbeitet in Vorwärts- und Rückwärtsbewegungen. In der Vorwärtsbewegung wird die Regressionsgleichung bewertet und in der Rückwärtsbewegung werden die Gewichte der einzelnen Regressoren optimiert. Das in diesem Abschnitt untersuchte Modul ist ebenfalls ein rückwärts gerichteter Optimierungsalgorithmus. Der Algorithmus funktioniert folgendermaßen: nachdem das Genetische Programm eine Gleichung entwickelt hat, setzt der Algorithmus zunächst am ersten Operanden an und gewichtet diesen sequentiell zwischen eins und minus eins mit Intervallschritten von 0.1. Somit entstehen insgesamt 21 neue Gleichungen inklusive der Originalfunktion. Anschließend werden die Gleichungen anhand des Fitneßkriteriums bewertet und die fitteste Gleichung wird gespeichert. Der Algorithmus verwendet jetzt die fitteste Gleichung und setzt am zweiten Operanden an und vollzieht dieselbe Prozedur. Das Modul arbeitet sich demnach von links nach rechts durch. Da dieser Vorgang sehr rechen- und daher zeitintensiv ist, wurde das Modul nur auf die beste Funktion angewandt. Die nachstehenden Abbildungen zeigen die Ergebnisse bei Anwenden des Moduls in jeder Generation und in jeder zwanzigsten. In Abbildung 2.21 sind die prozentualen Verbesserungen veranschaulicht, die durch den Algorithmus hervorgerufen worden sind. Gegenüber einem Modell mit einer Populationsgröße von 100 haben sich die Ergebnisse erkennbar verbessert (siehe Abbildung 2.20). Die Fitneßentwicklung beider Modelle sind sehr ähnlich, so daß es ausreicht, den Algorithmus in bestimmten Abständen und nicht in jeder Generation anzuwenden. Die prozentualen Verbesserungen sind bei diskreter Anwendung größer als bei kontinuierlicher Anwendung und besitzt den Vorteil einer erheblichen Zeitersparnis. Die Simulation bei kontinuierlicher Anwendung (100 Läufe x 200 Generationen bei einer Population von 100) dauerte ungefähr vier Tage. Die Ergebnisse indizieren,

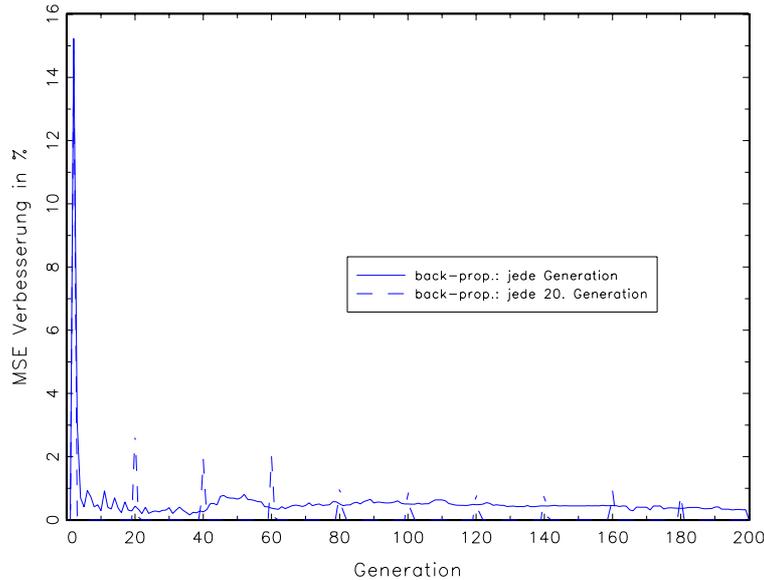


Abbildung 2.21: Verbesserung in Prozent bei Verwenden des Backpropagation Moduls

daß es sinnvoll erscheint, neue Optimierungsalgorithmen zu implementieren. Tatsächlich gibt es bereits Anwendungen, die die verschiedenen Disziplinen, die im ersten Kapitel kurz vorgestellt wurden, miteinander verbinden.

2.7.10 Zusammenfassung

Bei der Sensitivitätsanalyse wurde ein breites Spektrum an verschiedenen Parametern untersucht mit dem Ziel, die optimale Leistung des Genetischen Programms unter Berücksichtigung der benötigten Rechenzeit zu gewährleisten. Für die Analyse wurde der chaotische aber deterministische Prozeß von Tong zugrundegelegt. Die Ergebnisse sind teilweise recht überraschend. Die Variation der Funktionsmenge hat ergeben, daß eine problemorientierte Auswahl die Ergebnisse positiv beeinflussen kann. Es war diejenige Funktionsmenge am erfolgreichsten, die die Funktionsoperatoren des zu regressierenden Prozesses am ehesten widerspiegelte. Da einerseits die Funktionsmenge, die sämtliche Funktionsoperatoren umfaßt, nur marginal schlechtere Ergebnisse erzielte und andererseits die dynamischen Prozesse der zu schätzenden Finanzmarktmodelle unbekannt sind, werden alle Funktionsoperatoren berücksichtigt. Die Ergebnisse der Variation der Anteile genetischer Operationen bestätigen nicht die Daumenregel von Banzhaf et al., die 90% Crossover- und 10% Mutationsoperationen vorschlagen. Allerdings liegen die Ergebnisse sehr dicht beieinander. Im Folgenden wird ein Crossoveranteil von 60% und ein Mutationsanteil von 40% gewählt. Die Mutationen entfallen zur Hälfte auf Typ A und auf Typ B.

Die Untersuchungen der Knotenlänge der Anfangsgleichungen haben gezeigt, daß sich die Fitneßwerte mit Zunahme der Knotenlänge verschlechtern. Für die folgenden Untersuchungen wird eine Knotenlänge von fünf gewählt. Dagegen wird die Länge der Gleichungen auf 100 Knoten begrenzt und orientiert sich damit an der in der Literatur vorkommen-

den Größe. Die Ergebnisse fielen nur bei sehr kurzen Gleichungen deutlich schlechter aus. Die Selektionsmethode für genetische Operationen hat überraschende Ergebnisse im Hinblick auf die Tournament Methode hervorgebracht. Während Chen et al. die Tournament Methode im Vergleich zum proportionalen und uniformen Verfahren vorziehen, favorisieren die vorliegende Ergebnisse das proportionale Verfahren. Das Tournament Verfahren schneidet sogar schlechter als die uniforme Methode ab. Allerdings liegen die Ergebnisse erneut recht nahe zusammen. Für die weiteren Untersuchungen wird die proportionale Selektionsmethode angewendet. Bei der proportionalen Methode sollen wiederum die besten 100 Gleichungen für genetische Operationen herangezogen werden bei Zugrundelegung einer fitneßbasierten Auswahlwahrscheinlichkeit. Der Versuch, längere Funktionen bei der Fitneßbewertung nach der Methode der Informationskriterien zu bestrafen, verursacht einen stark negativen Effekt auf die MSE Entwicklung. Daher wird weiterhin das MSE als Hauptkriterium verwendet. Allerdings soll untersucht werden, ob die auf Informationskriterien basierten Modelle ebenfalls in der Lage sind, gute Prognosemodelle zu erstellen trotz eventuell schlechterer In-Sample Ergebnisse. Solche Modelle könnten im Fall von Overfitting Problemen nützlich sein.

Die Analyse der Building Blocks liefert ebenfalls interessante Ergebnisse. Bei der Untersuchung der Finanzmarktmodelle sollen die Fitneßänderungen verwendet werden, um zu überprüfen, ob ein Auswahlkriterium guter Prognosegleichungen gefunden werden kann. Das Implementieren eines Backpropagation Optimierungsverfahrens konnte die Fitneßwerte zusätzlich verbessern. Da sich die Ergebnisse zwischen einer stetigen und diskreten Anwendung des Verfahrens kaum unterscheiden, soll das diskrete Verfahren bei der Schätzung monetärer Größen ebenfalls eingesetzt werden.

Kapitel 3

Geldpolitik und die Einordnung nichtlinearer Schätzmethoden

Im weiteren Verlauf der Arbeit werden mit Hilfe Genetischen Programmierens Prognosemodelle für monetäre Größen erstellt, wobei zunächst anhand von Inflationsprognosen auf der Grundlage des Mishkin Ansatzes verschiedene Kriterien für eine optimale Auswahl von genetisch entwickelten Prognosegleichungen untersucht werden. Anschließend werden im Kapitel 5 unter Berücksichtigung der Ergebnisse des Kapitels 4 Wechselkursprognosen vorgenommen. Die Untersuchungen stützen sich dabei auf die Studien von Meese & Rogoff [89] sowie Plasmans et al. [104] bzw. Verkooijen et al. [127], die mit Hilfe künstlicher neuronaler Netze univariate Zeitreihen- wie auch multivariate Strukturmodelle geschätzt haben. Die nichtlinearen Regressionsmodelle des Genetischen Programms werden für zwei verschiedene Zeiträume geschätzt, wobei das Louvre Abkommen aus dem Jahr 1987 einen markanten Einschnitt in der geldpolitischen Landschaft darstellt und als Abgrenzungszeitpunkt für den zweiten Datensatz gesetzt wird. Aufgrund der Tatsache, daß mit Inflations- und Wechselkursprognosen makroökonomische Zeitreihen untersucht werden, die durch geldpolitische Maßnahmen in der Vergangenheit entscheidend beeinflußt wurden, ist es sinnvoll, einen Überblick über die geldpolitische Ausgestaltung der Deutschen Bundesbank sowie bedeutende internationale Maßnahmen für den Zeitraum von 1973 bis 2002 zu geben.

3.1 Ausgestaltung der Geldpolitik

Mit dem Beschluß des Gesetzes über die Deutsche Bundesbank vom 26. Juli 1957 wurden die Aufgaben der Deutschen Bundesbank festgelegt. Neben technischen Aufgaben wird die Bundesbank vor allem auf das wirtschaftspolitische Ziel der Währungssicherung verpflichtet. Die explizite Aufforderung zur Geldwertstabilität umfaßt sowohl den binnenwirtschaftlichen Aspekt der Preisniveaustabilität als auch die Stabilität des realen Wechselkurses im Außenverhältnis. Hierunter versteht die Bundesbank die Aufrechterhaltung der Kaufkraft

der heimischen Wahrung im Ausland¹. Die Geldpolitik folgte dabei verschiedenen wirtschaftstheoretischen berlegungen, die hufig mit wirtschaftstheoretischen Trennpunkten einhergingen. Die folgende kurze Skizzierung erstreckt sich ber den Zeitraum von 1973 bis 2002, so da sie mit der Zeitspanne der in spateren Untersuchungen verwendeten Daten bereinstimmt². Das einschneidendste Ereignis bildet das Ende des Bretton-Woods Systems am 19. Marz 1973. Es markiert das Ende fixer Wechselkurse und des Gold-Devisenstandards, wobei letzterer bereits mit dem Aussetzen der Einlosbarkeit des US-\$ in Gold am 15. August 1971 faktisch aufhorte zu existieren. In der internationalen Diskussion und der praktischen Umsetzung der Zentralbankpolitik fanden nunmehr die Lehre Friedmans groere Beachtung. Die Abkehr vom Keynesianismus zugunsten des Monetarismus sorgte fur eine geldmengenorientierte Strategie der Bundesbank. Diese zweistufige Strategie versucht durch den Einsatz geldpolitischer Instrumente, die Geldmenge direkt und mit Hilfe dieser Zwischenzielgroe die Inflationsrate indirekt zu kontrollieren. Allerdings wurde keine wie von Friedman vorgeschlagene regelgebundene Geldpolitik sondern vielmehr ein pragmatischer Weg gewahlt, der neben der Preisniveaustabilitat auch andere gesellschaftspolitische Ziele, vor allem einen stabilen Auenwert der Deutschen Mark, zu berucksichtigen versuchte. Die Phase zwischen 1973 und 1984 war am Devisenmarkt von sehr starken Wechselkursschwankungen gepragt, in der die Notenbanken, abgesehen von der amerikanischen Zentralbank, die Wechselkurse durch Devisenmarktinterventionen in Richtung ihres Fundamentalwertes lenken wollten. Dabei bildete der US-\$ das vorrangige Ziel der Zentralbankinterventionen. Die Wahrungsturbulenzen zusammen mit den beiden lpreiskrisen hatten einen erheblichen Einflu auf die Volatilitat der Inflationsraten. Als der US-\$ am 26. Februar 1985 mit 3,4690 DM/US-\$ seinen hochsten Wert verzeichnete, vereinbarten am 22. September 1985 die USA, Japan, Deutschland, England und Frankreich im sogenannten Plaza Agreement eine engere Zusammenarbeit, um eine 'geordnete' Abwertung des Dollars voranzutreiben. Der Dollar sollte durch koordinierte Devisenmarktinterventionen unter Mithilfe der Amerikaner in die Richtung seines Fundamentalwertes zuruckgefuhrt werden. Am 22. Februar 1987 traf sich die um Italien und Frankreich erweiterte Gruppe erneut, um diese erfolgreiche "hegemoniale Kooperation" im sogenannten Louvre-Accord Abkommen zu beenden. Diese Phase koordinierter Interventionen stellt einen zweiten bedeutenden Einschnitt dar, der durch eine separate Modellschatzung und Prognose berucksichtigt wird. Die Bundesbank hat wahrend der gesamten Zeit trotz zahlreicher Zielverfehlungen ihre Geldmengenpolitik fortgesetzt.

Im Zuge der Europaischen Wahrungsunion sind am 01. Januar 1999 zahlreiche Rechte und Instrumente auf die Europaische Zentralbank (EZB) ubertragen worden. Die EZB verfolgt dabei eine Zwei-Sulen-Strategie, die als Hauptelement einer stabilitatsorientierten Geldpolitik eine Preisstabilitat mit einer maximalen Inflationsrate von 2% gegenuber dem Vorjahr als Zielvorgabe festgeschrieben hat. Eine geldmengenorientierte Politik mit der Geldmenge M3 als Bezugsaggregat stellt eine der tragenden Sulen bei der strategischen Ausgestaltung dar, wobei dies aber unter Berucksichtigung der Aussichten fur die Preisentwicklung geschehen soll. Der Artikel 105 (1) des EZB-Vertrages weist daher Parallelen

¹Die Deutsche Bundesbank, 1993, S.13: Geldpolitische Aufgaben und Instrumente, Sonderdruck der Deutschen Bundesbank, Nr.7, 6.Auflage

²Die folgenden Ausfuhungen orientieren sich an dem Artikel von Richter [110], 1998.

zum bisherigen Ansatz der Deutschen Bundesbank auf. Allerdings wird die Preisniveaustabilität der EZB eindeutig als das vorrangige Ziel ausgewiesen. Die Verschiebung der Gewichtung zugunsten der Preisstabilität mag für den Betrachter unwesentlich erscheinen. Für die praktische Umsetzung der Zwei-Säulen Strategie sowie das im nächsten Abschnitt angeführte Inflation Targeting sind gute Inflationsprognosemodelle erforderlich. Letzteres verwendet - im Gegensatz zur zweistufigen indirekten Geldmengenpolitik - eine einstufige Endzielstrategie mit direkter Orientierung am Inflationsziel.

3.2 Inflation Targeting

Die geldpolitische Konzeption des "Inflation Targeting" wurde in vielen Staaten als Antwort auf hohe Inflationsraten und Unglaubwürdigkeit der heimischen Notenbank eingeführt, um das Vertrauen der Wirtschaftssubjekte in die eigene Währung zurückzugewinnen. Die Glaubwürdigkeit der Zentralbank sollte dabei durch mehrere Faktoren wieder hergestellt werden³: gesetzlich verankerte Unabhängigkeit der Zentralbank, Bekanntgabe einer konkreten Inflationsrate in einem bestimmten Zeitraum, Bekanntgabe der regelmäßig erstellten Inflationsprognosen, Sanktionen bei Nicht-Erreichen des erklärten Endziels und Publikationen beispielsweise der Sitzungsprotokolle der Zentralbankgremien. Durch die erhöhte Transparenz der Zentralbanken soll von vornherein unterbunden werden, daß die Wirtschaftssubjekte Inflationserwartungen aufbauen. Damit liegt der erste Schritt zur Vermeidung einer tatsächlichen Inflation für die Notenbank in der Beeinflussung der Inflationserwartungen. Folglich richtet die Zentralbank ihre Zinspolitik nicht nach einem Zwischenziel wie etwa der Geldmenge sondern direkt nach den Inflationserwartungen der Wirtschaftssubjekte aus. Bei dem konkreten Endziel der Inflationsrate kann es sich entweder um einen Punktwert oder um einen Zielkorridor handeln, in der sich die Inflationsrate bewegen soll. Die mögliche Alternative des Preisniveaus anstelle von Inflationsraten als Endzielgröße hätte den Nachteil, daß positive oder negative Schocks auf den Finanzmärkten kompensiert werden müßten, um wieder auf den Zielpfad des Preisniveaus zu gelangen. Um das vorgegebene Endziel zu erreichen, erstellen die Zentralbanken regelmäßig Analysen im Hinblick auf die zu erwartende Preissteigerung. Liegt die Inflationsprognose über dem Zielwert, ergreift die Zentralbank restriktivere geldpolitische Maßnahmen. Dabei folgt sie keiner regelgebundenen Politik, sondern sie muß sich lediglich dem festgelegten Handlungsrahmen unterwerfen. Der Nachteil des 'Inflation Targeting' im Vergleich zur Geldmengenorientierung besteht wegen des Fehlens der Zwischenzielsteuerung in den größeren Wirkungsverzögerungen. Letzteres wurde häufig als Begründung für eine geldmengen- oder wechselkursorientierte Geldpolitik angeführt. Um die Vorteile der beider Strategien miteinander zu kombinieren, hat der schwedische Ökonom Lars Svensson 1996 [117] vorgeschlagen, die Inflationsprognose als Zwischenziel zu verwenden. Sie stellt seiner Meinung nach aufgrund der engen Korreliertheit mit dem Endziel ein ideales und zudem transparentes Zwischenziel dar.

³Bofinger et al. [11], Illing, S.130 ff.[68]

3.3 Genetisch entwickelte Modelle - ein geldpolitisches Prognoseinstrument?

Um die Frage nach der Einordnung genetisch entwickelter Prognosemodelle im Rahmen des geldpolitischen Instrumentariums beantworten zu können, bedarf es der Definition der Begriffe Zwischenziel und Indikator⁴. Den Zentralbanken stehen verschiedene geldpolitische Instrumente wie die Geldbasis oder Geldmarktzinssätze zur Verfügung, die sie unmittelbar beeinflussen können. Die Änderung solcher Instrumente setzt eine Kettenreaktion auf andere ökonomische Größen bis zum Endziel der Inflationsrate in Gang, der als monetärer Transmissionsprozeß beschrieben wird. Bei einer geldmengenorientierten Strategie versucht die Zentralbank anhand der Instrumente die Zwischenzielgröße der Geldmenge so zu regulieren, daß das Endziel der Preis(niveau)stabilität erreicht wird. Die Strategie des "Inflation Targeting" hingegen arbeitet ohne Zwischenziele und verwendet vielmehr Indikatoren anhand derer das Endziel direkt kontrolliert und prognostiziert werden soll. Zwischenziele und Indikatoren müssen daher unterschiedlichen Anforderungen genügen. Um als Indikator in Frage zu kommen, muß ein genetisch entwickeltes Prognosemodell die Fähigkeit wenigstens eines der beiden folgenden Funktionen besitzen:

- Prognosefunktion
- Wirkungsfunktion

Mit den beiden Anforderungen ist das rechtzeitige Signalisieren von Inflationsgefahren oder die korrekte Wiedergabe von Wirkungsrichtungen bei Veränderungen geldpolitischer Instrumente gemeint. Die Anforderungen an ein Zwischenziel sind umfangreicher und umfassen neben der Prognose- *und* Wirkungsfunktion zusätzlich einen

- stabilen Zusammenhang zwischen Operationsziel-Zwischenziel-Endziel.

Demnach wird ein stabiler Zusammenhang zwischen dem unmittelbar kontrollierbaren Instrumentarium der Notenbank (Operationsziel) und dem Zwischenziel sowie Zwischenziel und Endziel gefordert. Die unterschiedlichen Anforderungen verdeutlichen, daß genetisch entwickelte Prognosemodelle bei der Ausgestaltung der Zentralbankpolitik die Aufgabe eines Indikators übernehmen könnten.

3.4 Methoden zur Gewinnung von Inflationserwartungen

Da die Inflationserwartungen der Wirtschaftssubjekte ein zentrales Element in der Ausgestaltung des "Inflation Targeting" darstellen, versuchen die Zentralbanken diese bei der Erstellung ihrer Inflationsprognosen zu integrieren. Für die Gewinnung von Inflationserwartungen gibt es zwei zentrale Methoden⁵: entweder werden die Inflationserwartungen direkt mit Hilfe von Umfragen ermittelt oder indirekt aus Finanzmarktdaten extrahiert.

⁴Bofinger et al. (1996), S.241 ff.

⁵Dieser Abschnitt lehnt sich an den Ausführungen von Harff (2000) [59] an.

3.4.1 Umfragen

Bei der Befragung von Wirtschaftssubjekten handelt es sich um ihre persönliche Einschätzung im Hinblick auf die weitere Entwicklung der Inflationsrate⁶. Folglich werden die Prognosen nicht anhand eines expliziten Modells oder strukturellen Annahmen getroffen und können über mehrere Perioden starken Schwankungen ausgesetzt sein.

Neben diesen Nachteilen sind es aber vor allem die technischen Eigenschaften, die eine Verwendung in geldpolitischen Prognosemodellen nur begrenzt ermöglichen. Der organisatorische, zeitliche und kostspielige Aufwand einer Umfrage ist zu groß, um jederzeit verfügbare Daten zu generieren. Verbunden mit dem zeitlichen Aufwand stellt sich das Problem, daß die Inflationserwartungen dazu neigen, dem tatsächlichen Inflationsprozeß nachzulaufen. Darüber hinaus existiert das Problem, daß die Ergebnisse aufgrund ausgewählter Umfrageteilnehmer nicht repräsentativ und die Messungen nur sehr grob ermittelt werden können. Neben diesen technischen Mängeln wird häufig kritisiert, daß die Antworten in einer Umfrage nicht dem tatsächlichen ökonomischen Handeln der Marktteilnehmer entsprechen müssen. Umfragen entspringen also keinem marktrelevanten Kalkül. Daher können sie nur bedingt die tatsächliche Inflation vorhersagen. Trotz ihrer zahlreichen Schwächen finden Inflationserwartungen aus Umfragen in der ökonomischen Literatur durchaus Beachtung⁷.

3.4.2 Finanzmarktdaten

Die den Umfrageerhebungen innewohnenden Nachteile existieren bei der Gewinnung finanzmarktbasierter Inflationserwartungen nicht. Sie sind jederzeit verfügbar, können relativ leicht gemessen werden und sind aufgrund der großen Anzahl von Marktteilnehmern durchaus repräsentativ. In der Literatur sind zwei Methoden anzutreffen, die anhand von Zinssätzen versuchen, Inflationserwartungen zu extrahieren.

Betrachtet man Abbildung 3.1, die den Verlauf deutscher Zinsentwicklungen von Null-Kupon-Anleihen mit verschiedenen Restlaufzeiten und der Inflationsrate widerspiegelt, so läßt der parallele Verlauf vermuten, daß Zinssätze Informationen über die zukünftige Entwicklung der Inflationsraten aufweisen. Die ältere Methode setzt bei der Zinsstruktur an und geht auf Frederic Mishkin zurück. Mit Hilfe der Erwartungstheorie wird begründet, daß die Zinsdifferenz zweier Anleihen unterschiedlicher Laufzeit bei Annahme konstanter Realzinsen Informationen hinsichtlich der zukünftigen Inflationsänderungsrate enthalten müßte:

$$\pi_t^m - \pi_t^n = \alpha_{m,n} + \beta_{m,n}[i_t^m - i_t^n] + \eta_t^{m,n} \quad (3.1)$$

mit π_t^m als durchschnittliche Inflationsrate⁸ von Periode t bis $t + m$, i_t^m als Nominalzins von Periode t bis $t + m$, $\alpha_{m,n}$ als Konstante für die durchschnittliche Differenz der entsprechenden Realzinsen und $\eta_t^{m,n}$ als Störgröße, inklusive aller Abweichungen der Realzinsen

⁶Ein Beispiel für Umfrageerhebungen sind die seit 1946 durchgeführten Livingston Daten. Sie gehen auf den Kolumnisten Joseph Livingston zurück und bilden die älteste Datenerhebung ihrer Art. Die Umfragen sammeln die Erwartungen von Ökonomen aus den Bereichen Industrie, Staat, Banken und Hochschulen.

⁷Vgl. zum Beispiel Keane/Runkle (1990) oder Reckwerth (1997).

⁸Der nominale und reale Zinssatz sowie die Inflationsrate werden als jährliche Durchschnittswerte berechnet: $\pi_t^m = (\ln(CPI)_t - \ln(CPI)_{t-m})/m$ mit $\ln(CPI)$ als logarithmierten Konsumentenpreisindex.

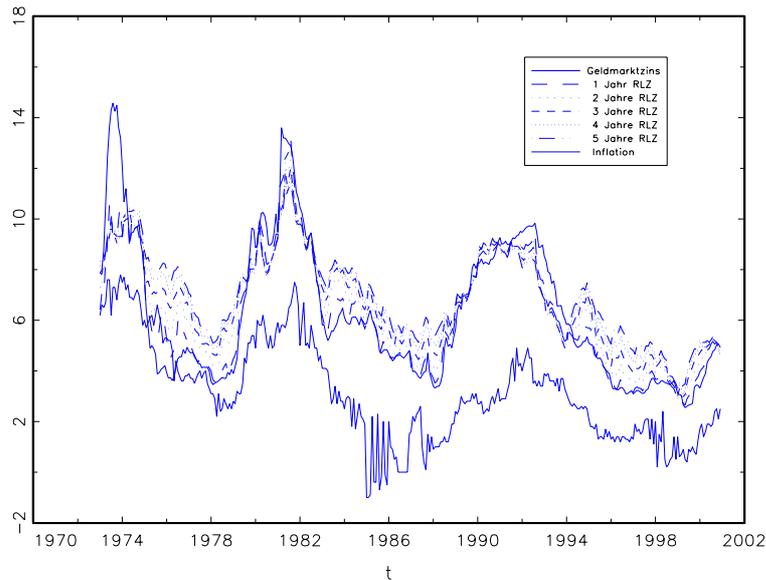


Abbildung 3.1: Zins- und Inflationsentwicklung von Januar 1973 bis Dezember 2000

von ihren Durchschnittsniveaus. Der Vorteil des Mishkin Ansatzes sind die über einen sehr langen Zeitraum hinweg zur Verfügung stehenden Zeitreihen. Der Nachteil besteht in der empirisch widerlegten Annahme konstanter Realzinsen.

Eine andere Methode, Inflationserwartungen aus Finanzmarktdaten zu extrahieren, verwendet indexierte Staatsanleihen. Diese Anleihen zeichnen sich durch die Kopplung der Nominalverzinsung an die Inflationsrate aus, so daß der Staat seine Schuldenlast durch eine überraschende Inflationierung nicht reduzieren kann. Obwohl das Konzept sehr alt ist, haben viele Staaten erst vor wenigen Jahren mit der Emission indexierter Staatsanleihen häufig im Zusammenhang mit der Einführung des "Inflation Targeting" begonnen. Aus der Zinsdifferenz nominaler Anleihen und indexierter Anleihen gleicher Laufzeit lassen sich Rückschlüsse auf die Inflationserwartungen der Marktteilnehmer ziehen. Da die Nominalzinsen neben einer realen Zins- und Inflationskomponente zusätzlich eine Risikoprämie für Volatilitätsschwankungen aufweisen, lassen sich die Erwartungen nicht unmittelbar aus dem direkten Vergleich ablesen. Erst bei Konstanz der Risikoprämie kann die Änderung der Zinsdifferenz Aufschluß über veränderte Inflationserwartungen geben. Für eine korrekte Gewinnung der Inflationserwartungen müssen die Anleihen zusätzlich identische Eigenschaften und einen hinreichend liquiden Handel aufweisen.

Die vorliegende Arbeit konzentriert sich auf den Ansatz von Mishkin und wendet die Methode auf deutsche Inflations- und Zinszeitreihen an.

Kapitel 4

Inflationsprognosemodelle

4.1 Mishkin Ansatz

Die Erwartungstheorie ist neben der Liquiditäts- und Marktsegmentierungstheorie eine der drei grundlegenden Erklärungsansätze für den Zusammenhang zwischen kurz- und langfristigen Zinssätzen¹. Mit ihr läßt sich sowohl ein normaler als auch ein inverser Verlauf der Zinsstruktur erklären. Kernpunkt der Theorie ist die Annahme, daß der durchschnittliche langfristige Zinssatz den in der betrachteten Periode erwarteten zukünftigen kurzfristigen Zinssätzen entspricht. Ein Anleger würde demnach den gleichen durchschnittlichen Zinssatz erwarten, unabhängig davon, ob er sein Geld fünf Jahre lang in einem langfristigen Wertpapier oder fünfmal hintereinander in kurzfristige Wertpapiere mit einjähriger Laufzeit anlegt. Der Zinsertrag einer langfristigen Anleihe muß im Arbitragegleichgewicht dem akkumulierten Ertrag kurzfristiger Anleihen entsprechen. Die Wertpapiere unterscheiden sich allein aufgrund der unterschiedlichen Laufzeit, wobei eine Risikoprämie für die langfristigen Wertpapiere in der ursprünglichen Erwartungstheorie ausgeschlossen wird. Somit werden die Wertpapiere unterschiedlicher Laufzeit als perfekte Substitute behandelt. Bezogen auf die obige Zinsstrukturgleichung bedeutet das folgendes: erwarten die Marktteilnehmer einen konstanten kurzfristigen Zinssatz in der Zukunft aufgrund unveränderter Inflationserwartungen, dann muß der Durchschnitt des langfristigen Zinssatzes in der gegenwärtigen Periode dem kurzfristigen Zinssatz entsprechen. Folglich würde die Zinsdifferenz in der Zinsstrukturgleichung des Mishkin Ansatzes null ergeben. Aufgrund der Annahme konstanter Realzinsen würden die Marktteilnehmer in diesem Fall auch keine Änderung der zukünftigen Inflationsrate erwarten. Wird dagegen ein Ansteigen der kurzfristigen Zinsen aufgrund höherer Inflationserwartungen vermutet, dann ist der langfristige Zinssatz (als Durchschnitt) größer als der gegenwärtige kurzfristige Zinssatz. Die positive Differenz in der Zinsstrukturgleichung prognostiziert wiederum eine Zunahme der zukünftigen Inflationsrate. Ein Ansteigen der Zinsstrukturkurve bildet daher ein Ansteigen der Inflationserwartung ab und umgekehrt. Je besser die Marktteilnehmer die zukünftige Inflationsentwicklung antizipieren, desto besser sollte der Mishkin Ansatz die Inflation vorhersagen können.

¹Vgl. auch Duwendag et al. [32], S.151 ff..

Gundlage zur Herleitung dieses Ansatzes ist das Fisher Theorem [41] aus dem Jahr 1930². Es besagt, daß sich der Nominalzinssatz i_t^m zum Zeitpunkt t und einer Restlaufzeit von m aus dem konstanten Realzinssatz r_t^m und der erwarteten Inflationsrate $E_t[\pi_t^m]$ zusammensetzt unter der Annahme eines effizienten Kapitalmarktes:

$$i_t^m = r_t^m + E[\pi_t^m] \quad (4.1)$$

Der von Fisher bereits 1916 erkannte Zusammenhang einer zeitweisen parallelen Entwicklung zwischen dem Preisniveau und dem nominalen Zinssatz wurde in empirischen Studien von Gibson 1923 und 1926 bestätigt. Vor allem die zeitliche Verzögerung der Zinsentwicklung gegenüber dem Preisniveau veranlaßte Fisher zu der Vermutung eines positiven Zusammenhangs zwischen den gegenwärtigen Inflationserwartungen der Marktteilnehmer und der zukünftigen Entwicklung des Nominalzinssatzes (Fisher Effekt)³. Jahrzehnte später wurde diese Hypothese von Fama (1975) [37] wieder aufgegriffen und konnte anhand amerikanischer Daten zunächst unterstützt werden. Fama's Ergebnisse lassen auf konstante (erwartete) Realzinsen im Zeitablauf schließen, so daß die Inflationserwartungen aus den Nominalzinsen extrahiert werden können. Allerdings mußte in nachfolgenden Arbeiten die Annahme konstanter Realzinsen wieder zurückgenommen werden unter anderem von Fama selbst⁴.

Da die erwartete Inflationsrate der Marktteilnehmer keine exakt meßbare Größe darstellt, verwendet Mishkin die ex-post realisierten Inflationsraten⁵. Diese entsprechen zum Zeitpunkt $t + m$ den erwarteten Inflationsraten zum Zeitpunkt t und einem Prognosefehler ε_t^m :

$$\pi_t^m = E[\pi_t^m] + \varepsilon_t^m \quad (4.2a)$$

$$\Rightarrow E[\pi_t^m] = \pi_t^m - \varepsilon_t^m \quad (4.2b)$$

Bei Annahme rationaler Erwartungen enthält ε_t^m keine systematischen Fehler $E_t[\varepsilon_t^m] = 0$, so daß man durch Einsetzen in die nach der erwarteten Inflationsrate aufgelösten Fisher Gleichung folgenden Ausdruck erhält:

$$E[\pi_t^m] = i_t^m - r_t^m \quad (4.3a)$$

$$\Rightarrow \pi_t^m = i_t^m - r_t^m + \varepsilon_t^m \quad (4.3b)$$

²Vgl. auch Bofinger et al. [11], S.330 ff. oder Duwendag et al. [32], S.20 ff..

³Angenommen ein Schuldner emittiert eine einjährige Anleihe in Höhe von 100,- Euro mit einer vereinbarten Zinszahlung von $r_0^1 = 5$ Prozent, dann werden die 100,- Euro am Ende der Laufzeit nur dann die gleiche Kaufkraft aufweisen, wenn in dieser Zeit keine Inflation $\pi_0^1 = 0$ aufgetreten ist. Um einen möglichen Realwertverlust aufgrund einer positiven Inflationsrate für den Gläubiger zu vermeiden, wird der Gläubiger die von ihm erwartete Inflationsrate $E_0[\pi_0^1]$ bei seiner Zinsforderung gegenüber dem Schuldner berücksichtigen. Somit enthält der nominale Zinssatz eine reale Komponente und die Inflationserwartungen der Marktteilnehmer. Der Kaufkraftausgleich für die nominalen Zinszahlungen $E[\pi_t^m] * r_t^m$ ist so gering, daß er bereits in der Fisher Gleichung vernachlässigt wurde und auch im Weiteren keine Beachtung findet. Vgl. auch Duwendag et al. [32], S.148.

⁴Nelson et al. [95] (1977), Mishkin [90] (1981), Fama und Gibbons [40] (1982)

⁵Zur Herleitung des Mishkin Ansatzes siehe auch Hansen [57] oder Jochum & Kirchgässner [70].

Entsprechendes gilt für einen mit Index n gekennzeichnete Inflationsrate π_t^n , so daß durch Subtrahieren schließlich die folgende Zinsstrukturgleichung entsteht:

$$\pi_t^m - \pi_t^n = i_t^m - i_t^n - r_t^m + r_t^n + \varepsilon_t^m - \varepsilon_t^n \quad (4.4)$$

Demnach wird die Veränderung der erwarteten zukünftigen Inflationsrate sowohl von der Steigung der nominalen als auch der realen Zinsstruktur bestimmt. Für die bekannte Prognosegleichung unterstellt Mishkin die problematische Annahme einer konstante Steigung der realen Zinsstruktur:

$$\pi_t^m - \pi_t^n = \alpha_{m,n} + \beta_{m,n}[i_t^m - i_t^n] + \eta_t^{m,n} \quad (4.5)$$

mit

$$\begin{aligned} \alpha_{m,n} &= \bar{r}^n - \bar{r}^m, \\ \beta_{m,n} &= 1, \\ \eta_t^{m,n} &= \varepsilon_t^m - \varepsilon_t^n - (u_t^m - u_t^n), \\ u_t^m &= r_t^m - \bar{r}^m, \\ u_t^n &= r_t^n - \bar{r}^n \end{aligned}$$

Während also bei der Fisher Gleichung die Inflationserwartung aus der Differenz zwischen Nominal- und (konstantem) Realzins extrahiert wird, versucht die Zinsstrukturgleichung die Inflationserwartungen aus der Differenz der Nominalzinsen zu gewinnen. Diese spiegelt dann die Änderung der Inflationserwartung und nicht das Niveau wider. Ein Ansteigen der Zinsstrukturkurve bildet daher ein Ansteigen der Inflationserwartung ab und umgekehrt. Die zugrunde gelegte Annahme einer konstanten Differenz bzw. Steigung der Realzinsstruktur schließt variable Realzinsen dabei nicht aus. Je besser die Marktteilnehmer die zukünftige Inflationsentwicklung antizipieren, desto besser sollte der Mishkin Ansatz die Inflation vorhersagen können.

4.2 Ergebnisse ausgewählter empirischer Studien

Zum zinsstrukturtheoretischen Ansatz existieren zahlreiche empirische Untersuchungen, die sich nicht nur in ihrer ökonometrischen Herangehensweise sondern auch in ihren Schlußfolgerungen zum Teil deutlich voneinander unterscheiden⁶. Der Überblick beginnt mit zwei Studien von Mishkin [91], [92] aus dem Jahr 1990 mit Untersuchungen amerikanischer Anleihen mit Restlaufzeiten von einem bis zwölf Monate und von einem bis fünf Jahre. Mishkin schätzt die Gleichung 4.5 mit Hilfe einer einfachen OLS Regression unter Zuhilfenahme ökonometrischer Korrekturverfahren inkorrektur Standardfehler. Die beiden Studien führen unter anderem zu dem Ergebnis, daß die nominale Zinsstruktur mit steigender Restlaufzeit auch einen steigenden Informationsgehalt hinsichtlich der zukünftigen Änderung der Inflationsrate enthält. Mit zunehmender Restlaufzeit steigen sowohl

⁶Die Übersicht lehnt sich an den Ausführungen von Kirchgässner & Jochum [70] sowie Hansen [57] an.

die β -Koeffizienten als auch die Bestimmtheitsmaße R^2 auf teilweise deutlich höheren Signifikanzniveaus. Die Arbeiten von Fama & Gibbons [40] (1982), Fama [38] (1984) und Fama [39] (1990) unterstützen unter Verwendung ebenfalls amerikanischer Daten diese Ergebnisse. Estrella & Mishkin (1997) [36] analysieren dagegen Zeitreihen aus Frankreich, Italien, Deutschland und Großbritannien und benutzen eine modifizierte Schätzgleichung. Für die nominale Zinsstruktur verwenden sie die Zinssätze einer 10-jährigen Anleihe sowie eines dreimonatigen Treasury Bill und betrachten Prognosehorizonte zwischen einem und zwanzig Quartale. Darüber hinaus werden zeitverzögerte Inflationsraten als endogene Variable in der Prognosegleichung berücksichtigt. Trotz dieser Änderungen kommen sie zum gleichen Ergebnis wie die Untersuchungen zuvor und bestätigen, daß die nominale Zinsstruktur auch in europäischen Ländern Informationen bezüglich der zukünftigen Inflationsänderungen besitzt.

Die bisher vorgestellten Studien benutzen ähnliche ökonometrische Methoden. Sie vernachlässigen aber die Überprüfung der verwendeten Zeitreihen auf Stationarität. Die in der Literatur bisher durchgeführten Untersuchungen zeigen, daß die Inflations- und Zinszeitreihen nichtstationär sind. OLS Schätzungen können daher „spurious regressions“ oder irreführenden Regressionen verursachen, worauf Granger & Newbold [54] in ihrem Aufsatz von 1974 hingewiesen haben. Das Bestimmtheitsmaß R^2 oder auch die t-Statistiken zur Beurteilung der Signifikanz können bei nichtstationären Variablen, die einem gemeinsamen Trend aufweisen, zu groß ausfallen und zu Fehlinterpretationen führen. Mishkin [93] (1992), Tzavalis und Wickens [126] (1996), Jochum und Kirchgässner [70] (1999) sowie Hansen [57] (2001) berücksichtigen diesen Aspekt in ihren Untersuchungen, wobei Mishkin die bivariate Kointegrationsmethode von Engle-Granger [34] (1987) und die anderen Autoren das multivariate Johansen Verfahren [71] (1988) bevorzugen. Mit diesen Kointegrationsmodellen lassen sich die langfristigen Beziehungen zwischen Nominalzinsen und Inflation analysieren wie auch die kurzfristigen Beziehungen durch Fehlerkorrekturmodelle ableiten. Jochum und Kirchgässner untersuchen den Mishkin Ansatz mit Monatsdaten für die Bundesrepublik Deutschland. Dabei testen sie die Variablen in bivariaten Testverfahren auf Kointegration - im Gegensatz zum simultanen alle Variablen einbeziehenden Johansen Verfahren. Die Zeitreihen weisen - mit einzelnen (testabhängigen) Abweichungen - nichtstationäre Eigenschaften auf. Gemäß diesem speziellen Testverfahren sind die Variablen zwar kointegriert, die empirisch langfristigen Beziehungen widersprechen jedoch den dem Mishkin Ansatz zugrunde liegenden theoretisch geforderten Abhängigkeiten. Daraus folgern Jochum und Kirchgässner, daß der Mishkin Ansatz zur Inflationsprognose für deutsche Daten ungeeignet ist.

Hansen kritisiert die ökonometrische Vorgehensweise von Jochum und Kirchgässner. Seiner Meinung nach kann man das vorliegende multivariate Problem nicht schrittweise in bivariate Tests zerlegen. Wie Jochum & Kirchgässner verwendet auch Hansen deutsche Monatsdaten für Nominalzinsen und Inflationsraten. Seine Stationaritätstests ergeben, daß die Zeitreihen integriert vom Grade eins und somit nichtstationär sind. Tests mit Hilfe des Johansen Verfahrens lassen auf insgesamt drei Kointegrationsvektoren schließen. Aus diesen langfristigen Beziehungen leitet Hansen Vektorfehlerkorrekturmodelle (VECM) ab, mit denen er die zukünftige Inflation prognostiziert. Es zeigt sich, daß die Inflationsraten

sowohl von der Zinsstruktur als auch von den Realzinsen abhängen. Da gemäß der Zinsstrukturtheorie eine konstante Risikoprämie für Wertpapiere angenommen wird, wurden die Kointegrationsbeziehungen inklusive einer Konstanten geschätzt. Im Vergleich zu den vorangegangenen Untersuchungen führt Hansen als einziger eine ex-ante Inflationsprognose durch. Im Gegensatz zu Jochum und Kirchgässner folgert Hansen, daß die Zinsstruktur zur Inflationsprognose für Deutschland sehr wohl geeignet ist und als Indikatorvariable für die Zentralbank nützlich sein kann.

Eine weitere Studie zu den deutschen Monatsdaten führten Carstensen & Hawallek [18] durch. Genau wie Hansen benutzen sie hierzu ein Vektorfehlerkorrekturmodell. Im Unterschied zu den bisherigen Untersuchungen schätzen sie ein Vier-Variablen System bestehend aus drei Zinszeitreihen unterschiedlicher Restlaufzeit sowie einer Inflationszeitreihe. Die Kointegrationsbeziehungen enthalten entweder einen modifizierten Zinsspread - Kombination aus drei anstatt der üblichen zwei Zinszeitreihen - und/oder eine reale Zinszeitreihe als zusätzlichen nichtstationären Faktor. Der Zinsspread aus dem ursprünglichen Mishkin Ansatz (Gleichung 4.5) wird durch die ermittelten Kointegrationsbeziehungen substituiert. Anschließend werden die modifizierten Gleichungen anhand einer OLS-Regression geschätzt und als Prognosegleichung für eine Out-of-Sample Vorhersage herangezogen. Die besten Prognoseergebnisse erzielen diejenigen Modelle, deren Kointegrationsbeziehungen sowohl die modifizierten Zinsspreads als auch die realen Zinsraten enthalten. Untersuchungen von Wolters (1998) [131] anhand deutscher Daten belegen, daß die Zinsspreads nichtstationär sind. Die ursprüngliche Mishkin Gleichung weist somit eine Fehlspezifikation auf. Der stationären Inflationsdifferenz auf der linken Seite steht der nichtstationäre Zinsspread gegenüber. Erst durch das Einbeziehen des zusätzlichen nichtstationären Faktors und der Kombination mit dem Zinsspread erreichen Carstensen & Hawallek eine stationäre Kointegrationsbeziehung. Die Autoren vermuten, daß dies der Grund für die negativen Ergebnisse bei Jochum & Kirchgässner ist. Zusammenfassend ergeben ihre Prognoseergebnisse, daß weder der herkömmliche noch der modifizierte Zinsspread in einer Out-of-Sample Studie besser abschneiden als der Random Walk. Werden allerdings neben dem modifizierten Zinsspread die realen Zinsraten berücksichtigt, dann wird der Random Walk geschlagen. Carstensen & Hawallek betonen aber, daß für bestimmte Perioden auch das beste Prognosemodell große Vorhersagefehler verursacht.

Insgesamt zeigen die Untersuchungen, daß die Ergebnisse alles andere als eindeutig sind. Die Unstimmigkeiten beginnen bereits bei der Feststellung, ob die Zeitreihen stationär oder nichtstationär sind. Die überwiegende Anzahl der Untersuchungen stützt zwar die Annahme der Nichtstationarität der Inflations- und Zinsvariablen, dennoch sind diese Ergebnisse stark abhängig vom gewählten Beobachtungszeitraum sowie der Wahl der Datenfrequenz. Obwohl sich die ökonometrischen Methoden weiterentwickelt haben, weisen selbst die neueren Verfahren bei den Studien von Jochum & Kirchgässner im Vergleich zu Hansen und Carstensen & Hawallek gegensätzliche Schlußfolgerungen auf.

4.3 Konstruktion von Vergleichsmodellen auf Basis deutscher Monatsdaten

Im Folgenden wird der Mishkin Ansatz auf deutsche Monatsdaten angewendet. Die hierfür benötigten Inflations- und Zinszeitreihen erstrecken sich über einen Zeitraum von Januar 1975 bis Dezember 2002, wobei für die Schätzung und der anschließenden Prognose der Modelle der Datensatz in eine In-Sample (Januar 1975 bis Dezember 1996) und eine Out-of-Sample Datenmenge (Januar 1997 bis Dezember 2002) geteilt wird. Für alle Modelle wird die zukünftige Inflationsrate in zwei Jahren berechnet, wobei als Gütemaß erneut die mittlere quadratische Fehlerabweichung (MSE) sowohl für die In-Sample wie auch der Out-of-Sample Schätzung verwendet wird. Die Inflationsraten berechnen sich als logarithmische Differenz aus dem Preisindex für Lebenshaltung aller privaten Haushalte und stammen vom Statistischen Bundesamt⁷. Als Zinszeitreihen wurden hypothetische Null-Kupon-Anleihen mit Restlaufzeiten zwischen einem und fünf Jahren gewählt, die mit Hilfe der Svensson Methode aus der Zinsstrukturkurve geschätzt worden sind und bei der Deutschen Bundesbank erhältlich sind⁸. Zum Bilden der Zinsdifferenzen wurde ein dreimonatiger Geldmarktzinssatz vom Zinssatz der Null-Kupon-Anleihen subtrahiert.

Um Aussagen über die Prognosegüte genetisch entwickelter Modelle treffen zu können, werden im Folgenden uni- und multivariate Vergleichsmodelle konstruiert, wobei die univariaten Vergleichsmodelle das triviale Modell eines Random Walks sowie ARIMA und ARFIMA Prozesse beinhalten. Beim multivariaten Modell handelt es sich um ein Vektorfehlerkorrekturmodell, das mit Hilfe der Johansen Methode geschätzt wird.

4.3.1 Test auf Stationarität

Die Bestimmung des Integrationsgrades der zugrunde liegenden Zeitreihen ist vor allem bei der Entwicklung des multivariaten Vektorfehlerkorrekturmodells relevant. Im Zusammenhang mit den vorgestellten Ergebnissen ausgewählter empirischer Studien wurde bereits auf die methodischen Mängel im Zuge der Nichtbeachtung nichtstationärer Zeitreihen aufmerksam gemacht. Bei Vorliegen nichtstationärer Zeitreihen können einfache OLS-Regressionen autokorrelierte Fehlerterme verursachen mit der Folge verzerrter Teststatistiken. Im weiteren wird die schwache Stationarität untersucht, die folgende Bedingungen an eine Zeitreihe stellt:

1. Konstanter Erwartungswert: $E(X_t) = \mu$
2. Konstante Varianz: $var(X_t) = E(X_t - \mu)^2 = \sigma^2$
3. Konstante Kovarianz: $cov(X_t, X_{t-k}) = E[(X_t - \mu)(X_{t-k} - \mu)]$

⁷Es handelt sich um den Preisindex für das frühere Bundesgebiet. Das Statistische Bundesamt hat gegen Ende der Dissertationsschrift die Daten rückwirkend für ungültig erklärt.

⁸Eine nähere Erläuterung der Svensson Methode findet sich im Anhang C.1.

Zeitreihe	Niveau		1.Differenzen	
	ADF-Wert ¹	Lags	ADF-Wert ¹	Lags
Null-Kupon-Anleihe 1 Jahr RLZ	-2.2176	3	-9.0403	2
Null-Kupon-Anleihe 2 Jahre RLZ	-2.2161	3	-14.305	2
Null-Kupon-Anleihe 3 Jahre RLZ	-1.9385	2	-14.418	1
Null-Kupon-Anleihe 4 Jahre RLZ	-1.7200	1	-14.358	0
Null-Kupon-Anleihe 5 Jahre RLZ	-1.6720	1	-14.454	0
Geldmarktzins 3 Monate RLZ	-2.1753	1	-11.373	0
Inflationrate	-2.5008	2	-19.318	1
durchschnittliche Inflationsrate	-2.2588	2	-17.972	1

Tabelle 4.1: Test auf Stationarität. RLZ = Restlaufzeit; Anzahl der Lags wurde mit AIC und BIC Kriterium bestimmt; ¹ kritische Werte des ADF Tests mit Konstante; Signifikanzniveau von 1%, 5% und 10% entsprechen -3.64, -2.95 und -2.61 (MacKinnon [84]).

Für den Hypothesentest wird der erweiterte Dickey-Fuller-Test (ADF) angewendet, dem folgende Testgleichung zugrunde liegt:

$$\Delta X_t = \alpha + \delta X_{t-1} + \sum_{j=2}^q \delta_j \Delta X_{t-j+1} + \epsilon_t \quad (4.6)$$

Der Test berücksichtigt aufgrund des Verlaufs der Zeitreihen eine Konstante α und ermittelt die Anzahl der benötigten Lags mit Hilfe der Akaike und Schwarz Informationskriterien. Folgende Hypothesen werden getestet:

$$\begin{aligned} H_0 : \delta &= 0, \text{ für Nichtstationarität, wenn } t_\delta > \tau \\ H_1 : \delta &< 0, \text{ für Stationarität, wenn } t_\delta < \tau \end{aligned} \quad (4.7)$$

Die erhaltenen ADF Teststatistiken t_δ werden mit den kritischen Werten τ von MacKinnon [84] verglichen. Die in Tabelle 4.1 belegen auf einem sehr hohen Signifikanzniveau, daß alle Zins- und Inflationszeitreihen einen Integrationsgrad von eins aufweisen.

4.3.2 Tests auf die Prognosegüte

Bereits im Kapitel 2 wurden zwei Tests vorgestellt, die man zu Aussagen bezüglich der Prognosegüte heranziehen kann. Während die Momentenschätzung des Jarque-Bera Tests Rückschlüsse über extreme fehlerhafte Prognosen sowie asymmetrische Richtungsprognosen zuläßt, kann anhand des Box-Ljung Autokorrelationstests hingegen offengelegt werden, ob im Regressionsmodell relevante ökonomische Variablen fehlen und eine Misspezifikation des Modells vorliegt. Im Folgenden sollen zwei weitere Testverfahren eingeführt werden⁹.

modifizierter Diebold-Mariano-Test

Die genetisch entwickelten Prognosemodelle sollen mit den Benchmarkmodellen verglichen werden. Um zu überprüfen, ob sich die Out-of-Sample Werte der Modelle statistisch signifikant voneinander unterscheiden, wird eine modifizierte Version des Diebold-Mariano

⁹Vgl. Schröder [113], S. 455 ff..

Testverfahrens [27] (1995) angewendet, der von Harvey, Leybourne und Newbold [60] (1997) konzipiert worden ist. Es wird die Nullhypothese auf Gleichheit der Prognosewerte getestet:

$$H_0 : E[f_{i,t}^2 - f_{RW,t}^2] = 0 \quad (4.8)$$

wobei $f_{i,t}^2$ die quadratischen Fehlerabweichungen von Prognosemodell i : $f_{i,t} = (Y_t - \hat{Y}_{i,t})^2$ und $f_{RW,t}^2$ die Prognosefehler des Random Walks beschreiben¹⁰. Die Teststatistik des modifizierten Diebold-Mariano-Tests (mDM) berechnet sich wie folgt:

$$mDM = K * \frac{\bar{d}}{\sqrt{\hat{V}(\bar{d})}} \quad (4.9)$$

mit Korrekturfaktor $K = (\frac{N+1-2*h+N^{-1}*h*(h-1)}{N})^{-1}$, N = Anzahl der Prognosen und h = Anzahl der Prognoseschritte. Die Variable d drückt die Differenz der quadrierten Prognosefehler $d_t = f_{i,t}^2 - f_{RW,t}^2$, \bar{d} den arithmetischen Mittelwert von d_t und $\hat{V}(\bar{d})$ die heteroskedastie- und autokorrelationskonsistent geschätzte Varianz von \bar{d} aus¹¹. Die Modifikation des Diebold & Mariano Tests spiegelt sich durch die zusätzliche Verwendung des Korrekturfaktors K wider. Der Korrekturfaktor ist notwendig, da der ursprüngliche Test von Diebold & Mariano bei kleinen Datenmengen wegen der verzerrten Varianzschätzung tendenziell zu große Teststatistiken erzeugt. Überschreitet die Teststatistik einen bestimmten kritischen Wert aus einer t-Verteilung mit $(N - 1)$ Freiheitsgraden, dann weichen die Prognosewerte nicht signifikant voneinander ab. Der Test soll ausschließlich bei den genetisch entwickelten Prognosemodellen angewendet werden, deren Fehlerabweichungen geringer ausfallen als beim besten Referenzmodell.

Richtungsprognosen

Neben den Punktprognosen sollen auch Richtungsprognosen erstellt werden. Im Unterschied zu den MSE Punktprognosen, wo das Modell neben der korrekten Richtungsänderung auch das Ausmaß der Änderung vorhersagen muß, beschränkt sich die Richtungsprognose auf die Prognose der Richtungsänderung. Aufgrund der weniger restriktiven Anforderung können Modelle mit schlechten Punktprognosen durchaus in der Lage sein, gute Richtungsprognosen zu erzeugen.

Die Methode testet, ob die Anzahl korrekter Richtungsprognosen systematisch von einer Zufallsprognose abweichen. Basis des Tests ist eine 2×2 -Kontingenztabelle, die in der Hauptdiagonalen (N_{11}, N_{22}) die Anzahl der korrekten Richtungsprognosen und in der Nebendiagonalen (N_{12}, N_{21}) die Anzahl der falschen Richtungsprognosen ausweist. Die Teststatistik berechnet sich demnach wie folgt:

$$RProg = \sum_{i,j=1}^2 \frac{(N_{ij} - \hat{E}_{ij})^2}{\hat{E}_{ij}} \quad (4.10)$$

¹⁰Der Test ist nicht auf das MSE Gütemaß beschränkt, sondern kann jede beliebige Kennzahl zur Bewertung der Prognosegüte verwenden.

¹¹Für die Schätzung der Varianz wird das Newey-West Verfahren verwendet.

	Tatsächliche Veränderung "steigt"	Tatsächliche Veränderung "fällt"	Randsumme
Prognose "steigt"	N_{11}	N_{12}	N_1
Prognose "fällt"	N_{21}	N_{22}	N_2
Randsumme	$N_{\cdot 1}$	$N_{\cdot 2}$	N (Gesamtsumme)

Tabelle 4.2: Aufbau der Kontingenztabelle beim χ^2 -Unabhängigkeitstest

mit $\hat{E}_{ij} = N_{i\cdot} * N_{\cdot j} / N$. Die Teststatistik ist χ^2 -verteilt mit einem Freiheitsgrad. Die Trefferquote $(N_{11} + N_{22})/N$ gibt den Anteil korrekter Richtungsprognosen an und nimmt maximal den Wert eins an. Ähnlich wie beim Jarque-Bera Test können auch hier Asymmetrien zugunsten steigender oder fallender Veränderungen berechnet werden.

4.3.3 Random Walk

Ein Vergleichsmodell, das in der Literatur üblicherweise herangezogen wird, ist der Random Walk. Er stellt einen autoregressiven Prozeß erster Ordnung AR(1) mit einem Einheitskoeffizienten $\delta = 1$ dar:

$$y_t = \delta y_{t-1} + \epsilon_t = y_{t-1} + \epsilon_t \quad (4.11)$$

mit $\epsilon_t \sim WN(0, \sigma^2)$. Durch rekursives Einsetzen und einem Startwert für y_0 kann der Random Walk umformuliert werden zu

$$y_t = y_0 + \sum_{i=1}^t \epsilon_i \quad (4.12)$$

Analog erhält man y_{t+h} . Die ersten beiden Momente zeigen einen konstanten Erwartungswert aber eine mit der Zeit steigende Varianz und somit die Nichtstationarität des Random Walks:

$$\begin{aligned} E[y_{t+h,t}] &= y_t + 0 \\ Var[y_{t+h,t}] &= h\sigma^2 \end{aligned}$$

Somit ist der aktuelle Wert y_t stets der optimale Prognosewert eines Random Walk Modells unabhängig vom gewählten Prognosehorizont h . Angewandt auf die vorliegende Problemstellung bedeutet dies, daß der beste Schätzwert für die zukünftige Inflationsrate in 24 Monaten stets die aktuelle Inflationsrate darstellt. Tabelle 4.3 zeigt die Ergebnisse für die In-Sample als auch Out-of-Sample Berechnungen.

4.3.4 Modellierung von Long Memory Effekten: ARIMA vs. ARFIMA

In den nächsten beiden Abschnitten sollen zwei univariate Zeitreihenmodelle erstellt werden, die eventuell existierende langfristige Abhängigkeiten auf unterschiedliche Weise einzufangen versuchen. Solche Long-Memory Effekte können über die Autokorrelationswerte

	In-Sample 1975:3-1996:12	Out-of-Sample 1997:1-2002:12
MSE	4.064	1.203
Richtungsprognose (Teststatistik)	51.341 (0.00016)	35.211 (3.097)

Tabelle 4.3: Punkt- und Richtungsprognose des Random Walks; Richtungsprognose in %

beschrieben werden. Bildet man die Autokorrelationen derartiger Zeitreihen für unterschiedliche Zeitverzögerungen ab, so zeichnen sich diese Zeitreihen vor allem durch eine langsamere Abfallgeschwindigkeit der Autokorrelationen als bei Zeitreihen aus, die keine Persistenzeigenschaften aufweisen. In diesem Fall nimmt der Einfluß zeitverzögerter Werte sehr schnell ab und man spricht von einem Short-Memory-Effekt. Die **AutoRegressiven (Fraktional) Integrierten Moving Average** Prozesse versuchen anhand zeitverzögerter Werte der zu regressierenden Variablen Y_t sowie zeitverzögerter Fehlerterme ε_t die Dynamik der Zielvariablen zu approximieren. Formal betrachtet, unterscheidet sich ein AR-FIMA Modell in seiner Ausgangsgleichung nicht von einem ARIMA Modell:

$$\Phi(B)(1 - B)^d Y_t = \Psi(B)\varepsilon_t, \quad (4.13)$$

mit ε_t als White Noise Prozeß, B als Backshift-Operator ($BY_t = Y_{t-1}$), Φ als AR- und Ψ als MA-Polynome. Die langfristigen Abhängigkeiten drücken sich in dem Term $(1 - B)^d$ aus mit d als ganzzahligem oder fraktionalem Differenzenparameter bei einem ARIMA bzw. ARFIMA Modell. Die Modellierung von Zeitreihen mit Long-Memory Effekten geht zurück auf Mandelbrot [85] (1968), wobei die Entdeckung solcher Eigenschaften allerdings dem Hydrologen Hurst(1951) [67] zugeschrieben werden muß, der die Persistenzeigenschaften bei Untersuchungen der jährlichen Wasserstandsänderungen des Nils beobachtete.

ARIMA(p,d,q)

Für die Konstruktion des ARIMA Modells (Box & Jenkins [13],1976) muß die Anzahl der Zeitverzögerungen p und q sowie der Differenzengrad d ermittelt werden¹². Die Modellspezifikationen für p und q werden mit Hilfe der Akaike und Schwarz Informationskriterien gewonnen und sind im Anhang C.2 aufgeführt. Der Differenzen- bzw. Integrationsgrad d wurde bereits durch Anwenden eines ADF Tests bestimmt und ist in Tabelle 4.1 abgebildet. Demnach besitzt die Inflationszeitreihe erst nach einer Differenzenbildung stationäre Eigenschaften und ist somit integriert vom Grade eins. Bei der Anzahl der Zeitverzögerungen p und q schlägt das BIC Informationskriterium im Vergleich zum AIC aufgrund einer stärkeren Bestrafung zusätzlicher Regressoren eine restriktivere Parametrisierung vor. Konkret favorisiert das BIC ein ARIMA(2,1,0) und das AIC ein ARIMA(2,1,3). Allerdings zeigen die im Anhang aufgeführten Tests auf die Schätzresiduen, daß beide ARIMA Modelle nicht korrekt spezifiziert sind. Beide Modelle lehnen die Nullhypothese normalverteilter

¹²Die Berechnung der Statistiken für die ARIMA und ARFIMA Modelle wurden mit entsprechenden GAUSS Modulen vorgenommen.

Schätzfehler ab und weisen signifikante Autokorrelationswerte für quadrierte Residuen auf. Da diese Fehlspezifikationen auch bei zusätzlichen Zeitverzögerungen nicht verschwinden, wird an den vorgeschlagenen Modellspezifikationen festgehalten und die In- und Out-of-Sample Schätzungen berechnet. Weil das ARIMA Modell aufgrund der Nichtstationarität der Inflationsraten deren Änderungsraten schätzt, müssen diese zunächst akkumuliert und zum aktuellen Niveau addiert werden, bevor das Bewertungskriterium des MSE berechnet werden kann¹³:

$$y_{t+h,t} = y_t + \Delta y_{t+1,t} + \dots + \Delta y_{t+h,t} \quad (4.14)$$

Tabelle 4.4 enthält die Ergebnisse der beiden Modelle. Es zeigt sich, daß das ARIMA(2,1,3) Modell bei Punktprognosen marginal bessere In-Sample wie auch Out-of-Sample Ergebnisse erzeugt als ein ARIMA(2,1,0) Modell. Im Vergleich zum Random Walk schneiden beide Modelle allerdings schlechter ab. Anders verhält es sich hingegen bei den Richtungsprognosen. Der Random Walk dominiert zwar nach wie vor die In-Sample Schätzungen, in Bezug auf die Richtungsvorhersagen erzielen die ARIMA Modelle aber deutlich bessere Trefferquoten mit leichten Vorteilen für das sparsamer parametrisierte Zeitreihenmodell.

	ARIMA(2,1,3)		ARIMA(2,1,0)	
	In-Sample 1975:3-1996:12	Out-of-Sample 1997:1-2002:12	In-Sample 1975:3-1996:12	Out-of-Sample 1997:1-2002:12
MSE	4.574	1.333	4.580	1.342
DoC (Teststatistik)	50.575 (0.069)	57.746 (0.837)	48.462 (0.077)	59.155 (1.185)

Tabelle 4.4: Punkt- und Richtungsprognosen eines ARIMA(2,1,3) und eines ARIMA(2,1,0) Modells

ARFIMA(p,d,q)

Die Autoren Granger und Joyeux [53] (1980) sowie Hosking [65] (1981) verallgemeinerten die univariaten ARIMA Modelle, in dem sie reelle Werte für den Integrationsgrad zuließen - im Gegensatz zu den ganzzahligen Werten bei den ARIMA Modellen. Für die Berechnung des fraktionalen Differenzenparameters d wird der Term $(1-L)^d$ mit Hilfe der Gamma-Funktion transformiert, um eine leichtere Handhabung der Binomialfunktionen zu erreichen. Mit der Gamma-Funktion

$$\Gamma(g) = \int_0^{\infty} x^{g-1} e^{-x} dx \quad (4.15)$$

kann die Fakultät für natürliche Zahlen folgendermaßen formuliert werden:

$$\Gamma(x+1) = x(x-1)(x-2)\dots 1 = n! = x\Gamma(x) \quad (4.16)$$

¹³Vgl. auch Diebold [28], S.331.

Somit läßt sich $(1 - B)^d$ wie folgt ausdrücken¹⁴:

$$(1 - B)^d = 1 - dB + \frac{d(d-1)}{2!}B^2 - \frac{d(d-1)(d-2)}{3!}B^3 + \dots \quad (4.17)$$

$$= \sum_{k=0}^{\infty} \frac{\Gamma(k-d)B^k}{\Gamma(-d)\Gamma(k+1)} \quad (4.18)$$

Für Werte von $0 < d < 1/2$ erhält man ein stationäres ARFIMA Modell mit Long-Memory Effekten, deren Kovarianz-Summe zu einem endlichen Wert konvergiert. Liegt der Gedächtnisparameter in dem Bereich $-1/2 < d < 0$ erhält man einen invertierbaren anti-persistenten und für $d > 1/2$ einen nichtstationären Prozeß, der in diesem Fall zunächst differenziert werden muß. In der Literatur werden mehrere Schätzverfahren für den Gedächtnisparameter d präsentiert¹⁵, wobei hier lediglich die heuristische Methode nach Hurst [67], die sogenannte R/S Statistik, und die semi-parametrische Methode nach Geweke und Porter-Hudak [48] erläutert und bei der späteren empirischen Anwendung auf Inflationszeitreihen berücksichtigt werden.

GPH - Schätzverfahren

Die Methode von Geweke und Porter-Hudak (GPH) ist ein semiparametrischer Ansatz und verwendet für die Parameterschätzung die Spektraldichte. Mit Hilfe des kurzen Frequenzbereichs ist es möglich, den Gedächtnisparameter d für stationäre Zeitreihen zu ermitteln:

$$f(\lambda) \sim c_f |\lambda|^{-2d}, \quad \text{mit } |\lambda| \rightarrow 0^+, \quad (4.19)$$

dabei beschreibt c_f eine positive Konstante und d erneut den Differenzenparameter. Die Spektraldichte geht gegen unendlich, wenn die Frequenzen $|\lambda|$ gegen null gehen. Für einen Wertebereich von $0 < d < 1/2$ liegt eine Zeitreihe mit Long-Memory Effekten vor. Durch Logarithmieren erhält man eine log-lineare Gleichungsfunktion:

$$\log f(\lambda) \sim \log c_f - 2d \cdot \log |\lambda|. \quad (4.20)$$

Gleichung 4.19 kann in Form eines Periodogramms ausgedrückt werden:

$$I(\lambda_k) = \frac{1}{2\pi N} \left| \sum_{t=1}^N Y_t \cdot e^{-it\lambda_k} \right|^2, \quad (4.21)$$

$I(\lambda_k)$ bezeichnet dabei das Periodogramm der Zeitreihe Y_t bei einer Frequenz λ_k . Jede der fundamentalen Frequenzen $\lambda_k = \frac{2\pi k}{N}$, mit $k = 0, 1, \dots, \frac{N-1}{2}$, stellt einen asymptotisch unverzerrten Schätzer der Spektraldichte $f(\lambda)$ dar. Gleichung (4.20) kann daher wie folgt approximiert werden:

$$\log I(\lambda_k) \sim \log c_f - 2d \log \lambda_k + \log \xi_k, \quad (4.22)$$

¹⁴Siehe auch Hassler [56] oder Brockwell et al. [16].

¹⁵Siehe hierzu auch Beran [7] und Baille [4].

Die Residuen $\xi_k = I(\lambda_k)/f(\lambda_k)$ sind unabhängig standardisierte exponentielle Zufallsvariablen mit $E(\log \xi_k) = C = -0.577$ als Euler'sche Konstante¹⁶. Durch eine OLS Schätzung anhand einer Vielzahl von logarithmierten Periodogrammen auf eine Konstante und einer nichtlinearen Frequenzfunktion kann schließlich der Gedächtnisparameter d ermittelt werden, wobei die nachstehende Re-Definition lediglich die Übersicht fördern soll.

$$\begin{aligned} y_k &= \log I(\lambda_k), \\ x_k &= \log \lambda_k, \\ \beta_0 &= \log c_f - c, \\ \beta_1 &= -2d, \end{aligned} \tag{4.23}$$

Dadurch ändert sich Gleichung 4.22 in eine Regressionsgleichung mit identisch und unabhängig verteilten Residuen $\epsilon_k = C + \log \xi_k$ mit Mittelwert Null:

$$y_k = \beta_0 + \beta_1 x_k + \epsilon_k. \tag{4.24}$$

Da β_0 und β_1 durch OLS geschätzt werden kann, entspricht die Parameterschätzung d dem Ausdruck $\hat{d} = \hat{\beta}_1/2$. Ob statistisch signifikante langfristige Abhängigkeiten vorliegen, kann anhand eines einfachen t-Tests von d überprüft werden. Es wird die Nullhypothese, daß kein Long-Memory Effekt ($d = 0$) vorliegt, getestet. Ein in der Literatur häufig erwähntes Problem ist die korrekte Wahl des Frequenzbereichs k , das üblicherweise eine Funktion der Datenmenge T ist. Ein häufig verwendeter Wert ist $k = T^{0.5}$. Diese Wahl könnte aber zu verzerrten Testergebnissen führen, so daß verschiedene Frequenzbereiche untersucht werden.

R/S Statistik

Eine alternative Schätzung langfristiger Abhängigkeiten bei Finanzmarktzeitreihen stellt die von Mandelbrot verfeinerte R/S -Statistik von Hurst dar. Die R/S -Statistik bildet die kumulierten Summen über die Abweichungen zwischen der Zeitreihe und seinem Mittelwert ab, wobei die Spanne der einzelnen kumulierten Summen mit der Laufvariablen u ansteigt:

$$R_N := \max_{1 \leq u \leq N} \left[\sum_{j=1}^u (Y_j - \bar{Y}_N) \right] - \min_{1 \leq u \leq N} \left[\sum_{j=1}^u (Y_j - \bar{Y}_N) \right], \tag{4.25}$$

mit \bar{Y}_N als Mittelwert. Durch Division anhand der Standardabweichung

$$S_N := \sqrt{\frac{1}{N} \sum_{j=1}^N (Y_j - \bar{Y}_N)^2} \tag{4.26}$$

erhält man schließlich die adjustierte R/S Statistik:

$$Q_N := \frac{R_N}{S_N} \tag{4.27}$$

¹⁶Siehe auch Beran [7], S. 95 ff.

Für große Zeitverzögerungen k verhält sich die adjustierte Spannweite proportional zu $k^{0,5+d}$:

$$R/S \sim \text{Konstante} * k^{0,5+d} \quad (4.28)$$

und

$$\ln(R/S) \approx \ln(\text{Konstante}) + (0,5 + d)\ln(k) \quad (4.29)$$

Liegen langfristige Abhängigkeiten vor, muß bei einer graphischen Analyse, bei der $\ln(Q)$ gegen $\ln(N)$ abgetragen wird, die Steigung $H > 1/2$ betragen. Die klassische R/S Statistik nach Hurst lehnt allerdings die Nullhypothese kurzfristiger Abhängigkeiten zu häufig ab, so daß die von Lo [81] entwickelte modifizierte Version in der Literatur bevorzugt und bei der vorliegenden Untersuchung ebenfalls angewendet wird.

Empirische Untersuchung

Im Folgenden wird die Schätzung des Gedächtnisparameters d nach der R/S und GPH Methode und die Anpassung eines ARFIMA Modells anhand der Inflationsraten vorgenommen, wobei sich der In-Sample Zeitraum von Januar 1975 bis Dezember 1996 mit insgesamt $n = 264$ Beobachtungswerten erstreckt.

Ergebnisse nach GPH-Schätzmethode

Die Tabelle 4.5 enthält die Schätzergebnisse des Long-Memory Parameters d nach der Methode von Geweke und Porter-Hudak, wobei die Log-Periodogramm-Regressionen für verschiedene Fourierfrequenzen durchgeführt wurden, um eventuelle Sensitivitäten offenzulegen. Demnach zeigt die Tabelle für alle Frequenzbereiche, daß der Integrationsgrad der Inflationszeitreihe signifikant größer als 0.5 ist. Wie bereits in Punkt 4.3.1 erwähnt, deuten die Ergebnisse auf eine nichtstationäre Inflationszeitreihe hin, die für die weitere Modellierung zunächst differenziert werden muß. Somit stützt die GPH Methode die Resultate der ADF Testergebnisse. Berechnet man den Gedächtnisparameter d auf der

	$d(0.50)$	$d(0.55)$	$d(0.60)$	$d(0.80)$
d	0.9559	1.1328	1.0246	0.9629
St.error	0.2099	0.1755	0.1464	0.0765
t-Statistik	4.5536	6.4529	7.00049	12.5937
p-Value	0.0000	0.0000	0.0000	0.0000

Tabelle 4.5: Gedächtnisparameter für deutsche Inflationsraten

Grundlage differenzierter Inflationsraten, so läßt der Frequenzbereich $d(0.5)$ mit einem Gedächtnisparameter von $d = 0.3799$ und einem akzeptablen Signifikanzniveau auf das Vorliegen langfristiger Abhängigkeiten schließen (Tabelle 4.6). Da der GPH Test für kurze Frequenzen $|\lambda| \rightarrow 0^+$ konzipiert wurde, sind die anderen Ergebnisse weniger interessant.

	$d(0.50)$	$d(0.55)$	$d(0.60)$	$d(0.80)$
d	0.3799	0.0559	0.1099	-0.1233
St.error	0.2099	0.1755	0.1464	0.0765
t-Statistik	1.8096	0.3184	0.75119	-1.6126
p-Value	0.0704	0.7502	0.4526	0.1068

Tabelle 4.6: Gedächtnisparameter für differenzierte Inflationsraten

	normale Inflationsraten
Hurst Exponent	0.251
Gedächtnisparameter	-0.249

Tabelle 4.7: Deutsche Inflationsraten: Gedächtnisparameter $d = H - 1/2$

Davon abgesehen weisen die t-Statistiken ein eindeutig nichtsignifikantes Niveau auf. Die Ergebnisse der R/S Statistik widersprechen den qualitativen Aussagen für normale und quadrierte Inflationsraten mit Werten von 0.251 und 0.017 zugunsten von Short-Memory Eigenschaften, während absolute Inflationsraten eine leichte Unterstützung für langfristige Abhängigkeiten geben.

Anpassung eines ARFIMA Modells

Die Schätzung eines ARFIMA Modells soll semi-parametrisch als auch parametrisch vorgenommen werden. Bei der ersten Methode wird das GPH-Schätzergebnis von $d(0.50)$ vorgegeben, während es beim parametrischen Verfahren simultan mit den anderen Parametern geschätzt wird. Die Anpassung eines ARFIMA Modells an die Inflationszeitreihe erfolgt mit Hilfe des Bayes-Schwarz-Kriteriums (BIC), das im Vergleich zum Akaike Informationskriterium zusätzliche Modellparameter in einem stärkerem Ausmaß bestraft. Tabelle 4.8 enthält die BIC Werte für verschiedene Lag-Ordnungen. Das minimale BIC Kriterium weist einen Wert von 486.48 auf, so daß sich eine Modellordnung von ARFIMA(3,d,1) einstellt. Die in Tabelle 4.9 aufgeführten Modellparameter signalisieren, daß weder in den AR- noch in den MA Komponenten Unit Roots vorliegen. Allerdings besitzen die beiden Terme ϕ_1 und ϕ_2 sehr kleine und zudem nichtsignifikante Koeffizienten, so daß das ARFIMA(3,d,1) mit einem ARFIMA(1,d,1) sowohl für $d = 0.3799$ als auch für ein parametrisch geschätztes Modell in der Out-of-Sample Schätzung verglichen werden soll¹⁷. Die Ergebnisse der Tabelle 4.10 zeigen, daß kein ARFIMA Modell in der Out-of-Sample Prognose besser abschneidet als der Random Walk, wobei das parametrisch geschätzte ARFIMA(1,d,1) Modell die beste MSE Punktprognose erzeugt. Daher soll es als Benchmarkmodell bei der weiteren Betrachtung dienen.

¹⁷Bei einer erneuten Parameterschätzung mit einer maximal möglichen Modellordnung von $p = 1$ und $q = 1$ favorisierte das Schwarz Informationskriterium die Modellspezifikation ARFIMA(1,d,1). Die Ergebnisse der Modellschätzung sind im Anhang C.3 auf Seite 153 aufgeführt.

	$q = 0$	$q = 1$	$q = 2$	$q = 3$	$q = 4$
$p = 0$	603.69	185431.85	488.63	492.15	488.64
$p = 1$	546.33	488.66	492.72	494.98	834.50
$p = 2$	488.58	489.32	492.16	489.52	499.72
$p = 3$	494.22	486.48*	492.01	497.58	498.02
$p = 4$	496.20	491.98	497.58	498.02	503.44

Tabelle 4.8: Anpassung der Modellordnung anhand des BIC Kriteriums. BIC Werte für verschiedene ARFIMA(p, d, q) Modellordnungen für differenzierte Inflationsraten, $n = 264$ und gegebenem GPH Schätzwert $d: d(0.5) = 0.3799$ (siehe auch Tabelle 4.6). Für die AR und MA Parameter wurden insgesamt bis zu 4 Lags betrachtet. Folgende Berechnungsgleichung liegt den BIC Werten zugrunde: $BIC = -2 * \loglik + (p + q + 1) * \ln(n - 1)$. *minimales BIC.

parameter	estimates	std.err.	t-statistic	p-value
ϕ_1	0.092158	0.080160	1.149670	0.251340
ϕ_2	0.014182	0.094258	0.150463	0.880516
ϕ_3	0.259622	0.062970	4.122910	0.000050
θ_1	0.863705	0.062612	13.794551	0.000000
d	0.379900	0.000000	.	.
σ	0.580006	0.018125	31.999575	0.000000

Tabelle 4.9: Parameterschätzungen eines ARFIMA(3, d , 1) Modells für normale, differenzierte Inflationsraten bei $n = 264$ und einem GPH Schätzwert von $d(0.5) = 0.3799$. Die optimale Anzahl der Lags für die AR und MA Parameter wurde anhand des BIC Kriteriums bei einer maximalen Betrachtung von 4 Lags berechnet.

4.3.5 VECM

Aufgrund der Nichtstationarität der Nominalzins- und Inflationszeitreihen sowie deren gemeinsamen Integrationsgrad soll als multivariates Vergleichsmodell ein Vektorfehlerkorrekturmodell (VECM) konstruiert werden¹⁸. Zur Schätzung des multivariaten Kointegrationsmodells wird das Johansen Verfahren unter Zuhilfenahme des Programms CATS angewendet¹⁹. Kointegriert sind Zeitreihen, wenn mindestens eine Linearkombination mit stationärer Eigenschaft existiert. Bei synchronem Verlauf der Zeitreihen ist diese Eigenschaft häufig erfüllt²⁰ und ein Blick auf die Kurvenverläufe der einzelnen Zeitreihen (Abbildung 3.1) läßt vermuten, daß zwischen den Inflations- und Zinszeitreihen Kointegrationsbeziehungen bestehen. Da die Zeitreihen keinen eindeutigen Trendverlauf aufweisen, soll in dem VECM keine Trendkomponente in den Daten verwendet werden. Des weiteren nimmt die

¹⁸Vgl. auch Hansen [57], Patterson [101] oder Hansen & Julius [58].

¹⁹Für die Schätzung der VECM Modelle werden Null-Kupon-Anleihen mit Restlaufzeiten zwischen einem und fünf Jahren und ein konstanter Inflationsprognosehorizont von zwei Jahren verwendet. Diese Vorgehensweise läßt sich mit den Ergebnissen von Estrella & Mishkin [36] rechtfertigen.

²⁰Vgl. Seddighi et al. [115], S.282.

Modell	MSE	Trefferquote
ARFIMA(3,0.3799,1)	2.2149	0.3472 (3.3350)
ARFIMA(1,0.3799,1)	1.8514	0.4225 (0.9336)
ARFIMA(1,d,1)	1.2941	0.3944 (1.5976)

Tabelle 4.10: Out-of-Sample Prognosen verschiedener ARFIMA Spezifikationen für den Zeitraum: 1997:1 bis 2002:12.

Zinsstrukturtheorie an, daß in den Zinsdifferenzen bei Wertpapieren mit längeren Laufzeiten eine konstante Risikoprämie enthalten ist. Daher wird im Modell eine Konstante in den Kointegrationsvektoren eingebaut, so daß sich insgesamt die Vorgehensweise an Hansen [57] orientiert. Gleichung 4.30 gibt die allgemeine Schreibweise des Vektorfehlerkorrekturmodells wieder :

$$\Delta y_t = \alpha \beta' y_{t-1} + \Gamma \Delta y_{t-1} + \epsilon_t \quad (4.30)$$

Die (3x5) β' -Matrix enthält drei theoretisch unabhängige Kointegrationsvektoren, die weiter unten hergeleitet werden. Multipliziert mit dem (5x1)-Vektor $y_{t-1} = (\pi^m, \pi^n, i^m, i^n, 1)$, der die einzelnen Variablen und eine Konstante umfaßt, bekommt man die kointegrierenden Linearkombinationen zwischen diesen Variablen. Störungen der langfristigen Gleichgewichtsbeziehungen ($\beta' y_{t-1} \neq 0$) werden durch die Anpassungskoeffizienten der (4x3) α -Matrix wieder ausgeglichen, wobei jeder Spaltenvektor mit einem der r Kointegrationsvektoren korrespondiert. Die kurzfristigen Anpassungsdynamiken werden durch die (4x4) Γ -Matrix bestimmt. Da die Terme $\alpha \beta' y_{t-1}$ aufgrund der Linearkombinationen und $\Gamma \Delta y_{t-1}$ wegen der Differenzen stationär sind, ist auch das ganze System 4.30 stationär. Im Unterschied zu einem Vektorautoregressionsmodell (VAR) in Differenzen besitzt das VECM wegen des Terms $\alpha \beta' y_{t-1}$ den Informationsgehalt der langfristigen Beziehungen. Bevor anhand des Johansen Verfahrens die empirische Anzahl der Kointegrationsvektoren berechnet wird, soll zunächst die theoretisch zu erwartende Anzahl ermittelt werden, die sich aus der formalen Herleitung des Mishkin Ansatzes ergibt. Wie unter Punkt 4.1 bereits erwähnt, ist das Fisher Theorem unter Annahme rationaler Erwartungen Ausgangspunkt des Mishkin Ansatzes. Löst man die Gleichung 4.3b (S.64) nach dem Realzins auf, dann erhält man Gleichung 4.31:

$$r_t^m = i_t^m - \pi_t^m + \epsilon_t^m \quad (4.31)$$

Die Differenz der beiden nichtstationären Inflations- und Nominalzinszeitreihen läßt auf stationäre Realzinsen schließen, so daß diese Differenz eine Kointegrationsbeziehung darstellt. Aufgrund der Verwendung einer m - und n -periodigen Realzinsgleichung zur Herleitung des Mishkin Ansatzes liegen somit die ersten beiden theoretischen Gleichgewichtsbeziehungen vor. Als nächstes wird die Prognosegleichung betrachtet:

$$\pi_t^m - \pi_t^n = i_t^m - i_t^n + \epsilon_t \quad (4.32)$$

wobei $\epsilon_t = (r_t^n - r_t^m) - (u_t^n - u_t^m)$ entspricht. Für gegebene stationäre Realzinsen müssen die Inflations- und Nominalzinsdifferenzen $\pi_t^m - \pi_t^n$ bzw. $i_t^m - i_t^n$ ebenfalls stationär sein,

so daß der Mishkin Ansatz eine dritte und vierte Kointegrationsbeziehung erwarten läßt. Die Kointegrationsmatrix β sieht daher folgendermaßen aus:

$$\beta y_{t-1} = \begin{pmatrix} -1 & 0 & 1 & 0 & 1 \\ 0 & -1 & 0 & 1 & 1 \\ 0 & 0 & 1 & -1 & 1 \\ 1 & -1 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} \pi_{t-1}^m \\ \pi_{t-1}^n \\ i_{t-1}^m \\ i_{t-1}^n \\ 1 \end{pmatrix} \quad (4.33)$$

Die β -Matrix weist einen (Kointegrations-)Rang von drei auf, da der dritte und vierte Kointegrationsvektor nicht unabhängig sind. Den dritten Zeilenvektor erhält man beispielsweise durch eine lineare Transformation des ersten Zeilenvektors, in dem man vom ersten Zeilenvektor den zweiten subtrahiert und den vierten addiert. Unter Berücksichtigung der ersten drei unabhängigen Kointegrationsbeziehungen kann das System folgendermaßen geschrieben werden:

$$\begin{pmatrix} \Delta \pi^m \\ \Delta \pi^n \\ \Delta i^m \\ \Delta i^n \end{pmatrix} = \alpha \begin{pmatrix} -1 & 0 & 1 & 0 & 1 \\ 0 & -1 & 0 & 1 & 1 \\ 0 & 0 & 1 & -1 & 1 \end{pmatrix} \begin{pmatrix} \pi_{t-1}^m \\ \pi_{t-1}^n \\ i_{t-1}^m \\ i_{t-1}^n \\ 1 \end{pmatrix} + \Gamma \Delta y_{t-1} + \epsilon_t \quad (4.34)$$

Im Folgenden werden die Vektorfehlerkorrekturmodelle für verschiedene Zinsspreads bei konstanter durchschnittlicher zweijähriger Inflationsrate hergeleitet. Die Modellanpassung beginnt mit der Bestimmung der Laglänge p und somit mit der Anzahl der zeitverzögerten differenzierten Variablen Δy_{t-p} . Dies wird erneut anhand der Schwarz-, des Akaike- und Hannan-Quinn Kriterien vorgenommen. Liegen keine eindeutigen Aussagen vor, wird die kleinste Laglänge gewählt. Begründen kann man diese Vorgehensweise anhand der Simulationsergebnisse, die Lütkepohl bei Untersuchungen verschiedener Informationskriterien gewonnen hat²¹. Demnach unterschätzt das Schwarz Kriterium zwar die korrekte Laglänge am häufigsten, aber die sparsamere Parametrisierung führte zu den besten Prognoseergebnissen. Wie man aus den nachstehenden Gleichungen erkennen kann, bestraft das Schwarz Kriterium eine Zunahme der Freiheitsgrade am stärksten, so daß letztendlich das Schwarz Kriterium das entscheidende Kriterium ist. Für die Berechnung der Kriterien wird ein unrestringiertes VAR-Modell mit unterschiedlicher Laglänge geschätzt, wobei lediglich die In-Sample Daten bis Dezember 1996 verwendet werden. Unter Verwendung der jeweiligen Varianz-Kovarianzmatrix der Residuen $\hat{\Sigma}_u = u_t u_t' / n$ berechnen sich die Informationskriterien wie folgt:

$$\begin{aligned} AIC &= \ln|\hat{\Sigma}_u| + \frac{2m^2q}{n} \\ SIC &= \ln|\hat{\Sigma}_u| + \frac{m^2q}{n} \ln(n) \\ HQ &= \ln|\hat{\Sigma}_u| + \frac{2\ln(\ln(n))}{n} m^2q \end{aligned}$$

²¹Siehe Lütkepohl [82], S. 138

Die Tabelle 4.11 zeigt die optimalen Laglängen der einzelnen Informationskriterien. Das Schwarz Kriterium schätzt die optimale Laglänge bei allen Modellen auf zwei Lags, wohingegen das Akaike und Hannan-Quinn Kriterium drei Zeitverzögerungen signalisieren. Wendet man den gesamten Datensatz bis Dezember 2002 an, so tendiert das Hannan-Quinn Kriterium ebenfalls zu zwei Laglängen. Aufgrund dieser Ergebnisse wird eine Laglänge von zwei gewählt, so daß zur Schätzung des VECM ein VAR-Modell der Ordnung zwei zugrunde liegt. Da sich das VECM durch Subtrahieren des Terms $y_{t-1} = Iy_{t-1}$ auf bei-

$\pi^2 i^m$	AIC	SIC	HQ
$\pi^2 i^1$	3	2	3
$\pi^2 i^2$	3	2	3
$\pi^2 i^3$	3	2	3
$\pi^2 i^4$	3	2	3
$\pi^2 i^5$	3	2	3

Tabelle 4.11: Auswahl der Laglänge für das VECM anhand deutscher Daten 1973-1996.

den Seiten des VAR Modells formal herleitet mit I als Einheitsmatrix, reduziert sich die Laglänge des VECM auf eins. Damit stellt das System 4.30 bereits das vollständige zu schätzende Modell dar. Nach der Bestimmung der Laglänge muß als nächstes anhand des unrestringierten VAR(2) Modells der Kointegrationsrang der Matrix $\Pi (= \alpha\beta')$ ermittelt werden. Hierzu berechnet man die Eigenwerte der Matrix Π und überprüft mit Hilfe zweier Testverfahren, ob und wieviele Eigenwerte signifikant von null verschieden sind. Die Eigenwerte berechnen sich anhand des charakteristischen Polynoms $\Pi - \lambda_k I$, wobei mit λ_k die k resultierenden und geordneten Eigenwerte $\lambda_1 > \dots > \lambda_k$ bezeichnet werden²². Johansen schlägt die *trace*- und λ_{max} Teststatistiken für die Ermittlung des Kointegrationsrangs vor. Bezeichnet man mit $H(r_0)$ die Hypothese, daß der Rang r von Π kleiner gleich r_0 ist, dann testet das *trace*-Verfahren diese Nullhypothese gegen die alternative Hypothese $H(k) : r \geq r_0 + 1$, wobei k die Anzahl der endogenen Variablen darstellt:

$$trace(r_0|k) = -T \sum_{i=r_0+1}^k \ln(1 - \hat{\lambda}_i) \quad (4.37)$$

²²Genau genommen, berechnen sich die Eigenwerte aus der folgenden Gleichung:

$$|\lambda \hat{y}'_{t-1} \hat{y}_{t-1} - \hat{y}_{t-1} \Delta \hat{y} (\Delta \hat{y}' \Delta \hat{y})^{-1} \Delta \hat{y} \hat{y}'_{t-1}| = 0 \quad (4.35)$$

wobei $\Delta \hat{y} = [I_T - \Delta y^{*'} (\Delta y^{*'} \Delta y^*)^{-1} \Delta y^*] \Delta y$ und $\hat{y}_{t-1} = [I_T - \Delta y^{*'} (\Delta y^{*'} \Delta y^*)^{-1} \Delta y^*] y_{t-1}$ die Residuen einer Regression von Δy bzw. y_{t-1} auf die verzögerten differenzierten Variablen $\Delta y^* = (\Delta y_{t-1}, \Delta y_{t-2}, \dots, \Delta y_{t-p})$ darstellen. Die Eigenwerte der Gleichung 4.35 können äquivalent aus der Gleichung

$$|\lambda I - P(\hat{y}_{t-1} - \Delta \hat{y} (\Delta \hat{y}' \Delta \hat{y})^{-1} \Delta \hat{y}' \hat{y}_{t-1}) P'| = 0 \quad (4.36)$$

berechnet werden. $P = W^{-1}$ beschreibt dabei die Inverse aus der Choleski-Zerlegung von $\hat{y}'_{t-1} \hat{y}_{t-1} = WW'$. Man kann zeigen, daß $\sqrt{\lambda_i}$ die singularen Werte der Matrix $\hat{B}(1)$ im Hinblick auf $\hat{y}'_{t-1} \hat{y}_{t-1}$ und $(\Delta \hat{y}' \Delta \hat{y})^{-1}$ sind.

		Rangtest		χ^2 -Test auf Restriktionen	Kointegrationskonstante
1	2	3	4	5	6
π^m, i^m	r	λ_{max}	<i>trace</i>		
π^2, i^1	0	39.69*	75.44*	$\chi^2(3) = 1.42$	-3.174 (0.500)
	1	15.93*	35.74*	p-Wert = 0.70	-3.089 (0.544)
	2	12.83*	19.81	LM(1) = 0.00	-0.017 (0.234)
	3	6.98	6.98	LM(4) = 0.04	
π^2, i^2	0	40.53*	75.38*	$\chi^2(3) = 1.73$	-3.494 (0.431)
	1	14.77	34.85*	p-Wert = 0.63	-3.115 (0.548)
	2	12.71	20.07*	LM(1) = 0.00	-0.314 (0.328)
	3	7.36	7.36	LM(4) = 0.90	
π^2, i^3	0	40.25*	74.44*	$\chi^2(3) = 2.29$	-3.761 (0.402)
	1	13.92	34.19	p-Wert = 0.51	-3.142 (0.547)
	2	12.68	20.28	LM(1) = 0.00	-0.557 (0.401)
	3	7.60	7.60	LM(4) = 0.97	
π^2, i^4	0	39.90*	74.39*	$\chi^2(3) = 2.88$	-3.984 (0.379)
	1	14.56	34.49	p-Wert = 0.41	-3.160 (0.544)
	2	12.43	19.93	LM(1) = 0.00	-0.758 (0.453)
	3	7.49	7.49	LM(4) = 0.99	
π^2, i^5	0	39.69*	74.84*	$\chi^2(3) = 3.34$	-4.153 (0.361)
	1	15.43*	35.15*	p-Wert = 0.34	-3.171 (0.540)
	2	12.42*	19.73	LM(1) = 0.00	-0.915 (0.489)
	3	7.31	7.31	LM(4) = 0.99	

Tabelle 4.12: *Wert ist auf dem 5% Niveau signifikant. Die 5%-kritischen Werte für (r=0,1,2,3) sind für $\lambda_{max} = (18.69; 14.80; 10.86; 9.13)$ und für die trace -Statistik $trace = (53.42; 34.80; 19.99; 9.13)$.

T entspricht der Anzahl der Beobachtungswerte und λ_i den geordneten Eigenwerten. Ist die *trace*-Teststatistik größer als der entsprechende kritische Wert, dann wird $H(r_0)$ verworfen und $H(r_0 + 1)$ getestet. Diese Methode wird solange sequentiell fortgeführt, bis die Nullhypothese angenommen wird. Ähnlich verhält es sich mit dem λ_{max} -Testverfahren. Dessen Teststatistik sich folgendermaßen berechnet:

$$\lambda_{max}(r_0) = -T \ln(1 - \hat{\lambda}_i) \quad (4.38)$$

Der Unterschied zum *trace*-Test besteht darin, daß die Nullhypothese $H(r_0)$ gegen die explizite Alternativ-Hypothese $H(r_0 + 1)$ getestet wird. In Tabelle 4.12 sind die Ergebnisse der Testverfahren für alle untersuchten Modelle aufgelistet. Für die empirischen Daten zeigt sich, daß einige Systeme einen Kointegrationsrang von lediglich eins aufweisen. Allerdings sind die Teststatistiken nur marginal kleiner als der entsprechende kritische Wert. Erst beim Testen auf drei Kointegrationsränge weicht die Teststatistik deutlich vom kritischen Wert ab, so daß man durchaus einen Kointegrationsrang von drei wählen kann.

Unterstützt wird diese Annahme durch das Testen der Kointegrationsvektoren auf Nichtstationarität. In Tabelle 4.13 sind die ADF t -Werte der Kointegrationsvektoren sämtlicher Systeme aufgelistet. Mit Ausnahme der Realzinsbeziehungen der Systeme mit Null-Kupon-Anleihen mit Restlaufzeiten von einem und fünf Jahren weisen alle Gleichgewichtsbeziehungen auf einem Signifikanzniveau von 5% stationäres Verhalten auf. Bei einem Signifikanzniveau von 10% kann die Stationarität der beiden Ausnahmen jedoch ebenfalls nicht mehr abgelehnt werden. Auch die Eigenwerte der Companion Matrix (Tabelle 4.14) des

Kointegrationsvektor	ADF-Testwert	Laglänge
$i^0 - \pi^0 - c_{11}$	-3.1709	5
$i^1 - \pi^2 - c_{22}$	-2.9463	0
$i^1 - i^0 - c_{21}$	-5.0871	3
$i^0 - \pi^0 - c_{11}$	-3.1103	3
$i^2 - \pi^2 - c_{22}$	-3.1445	0
$i^2 - i^0 - c_{21}$	-3.2413	0
$i^0 - \pi^0 - c_{11}$	-3.1614	5
$i^3 - \pi^2 - c_{22}$	-3.1763	0
$i^3 - i^0 - c_{21}$	-4.5922	4
$i^0 - \pi^0 - c_{11}$	-3.1614	5
$i^4 - \pi^2 - c_{22}$	-3.2026	0
$i^4 - i^0 - c_{21}$	-4.3181	4
$i^0 - \pi^0 - c_{11}$	-3.1103	3
$i^5 - \pi^2 - c_{22}$	-2.7855	1
$i^5 - i^0 - c_{21}$	-4.1470	4

Tabelle 4.13: Test der einzelnen Kointegrationsvektoren auf Stationarität.

restringierten VAR Modells lassen auf drei Gleichgewichtsbeziehungen schließen. Der Kointegrationsrang von drei impliziert bei einem System mit vier endogenen Variablen genau einen nichtstationären explosiven Prozeß bzw. das Vorliegen einer Unit Root. Da CATS die Eigenwerte des charakteristischen Polynoms berechnet, müssen die Eigenwerte kleiner eins sein, um stationäre Linearkombinationen anzuzeigen. In allen untersuchten Modellen liegen die ersten beiden Eigenwerte relativ dicht an einer Unit Root und weisen somit einen nichtstationären explosiven Prozeß auf. Der dritte Eigenwert liegt etwas deutlicher unterhalb einer Unit Root, so daß die Companion Matrix ebenfalls drei Kointegrationsvektoren signalisiert. Bei der Annahme von drei Kointegrationsvektoren benötigt man mindestens zwei Restriktionen pro Kointegrationsvektor, um diese eindeutig bestimmen zu können. Da aber drei Restriktionen pro Vektor auferlegt werden, sind diese Restriktionen überidentifiziert. Anhand eines χ^2 -Tests in Abhängigkeit von der Anzahl der überidentifizierenden Restriktionen wird getestet, ob die Restriktionen der Zinsspreads und Realzinsen verworfen werden müssen. Der kritische Wert für $\chi^2(3)$ auf einem Signifikanzniveau von 5% beträgt 7.815. Die Teststatistiken in Tabelle 4.12 sind in allen Modellen deutlich kleiner, so daß die Restriktionen nicht abgelehnt werden können. Dies wird auch durch die marginalen Signifikanzniveaus (p-Werte) bestätigt, die allesamt größer als 5% sind.

Real	Complex	Modulus
0.9443	0.0383	0.945
0.9443	-0.0383	0.945
0.8936	0	0.8936
-0.3583	-0.0209	0.3589
-0.3583	0.0209	0.3589
0.30060	0	0.3060
0.1611	0	0.1611

Tabelle 4.14: Der Kointegrationsrang von drei impliziert eine Unit Root. Erst der Wert 0.8936 ist etwas deutlicher vom Wert eins entfernt, so daß auch die Companion Matrix eher einen Kointegrationsrang von drei unterstützt.

Um die Restriktionen akzeptieren zu können, muß untersucht werden, ob zum einen die geschätzten Kointegrations-Koeffizienten und zum anderen die Anpassungsparameter mit den theoretisch geforderten übereinstimmen. Tabelle 4.15 enthält die Anpassungsparameter der Kointegrationsvektoren zusammen mit den jeweiligen t-Werten. Im Vergleich zu den Ergebnissen von Hansen [57] existieren deutliche Unterschiede, die dadurch begründet werden können, daß die Zeitdifferenz $m - n$ der Inflationsraten $\pi_t^m - \pi_t^n$ bei Hansen stets mit dem Zinsspread $i_t^m - i_t^n$ übereinstimmt²³. Dagegen basieren die vorliegenden Untersuchungsergebnisse auf der Betrachtung einer konstanten zweijährigen Zeitdifferenz für die Inflationsrate und Variation des Zinsspreads in Bezug auf die Differenz der Restlaufzeiten. Ebenso wie bei Hansen nehmen die Anpassungskoeffizienten mit Zunahme des Zinsspreads ab. Weiterhin zeigen die Ergebnisse, daß die Zinsspreads und die kurzfristigen Realzinsen den kurzfristigen Nominalzinssatz signifikant bestimmen. Die aktuelle Änderungsrate der Inflation wird hingegen von allen Kointegrationsbeziehungen unterstützt. Auch wenn die Ergebnisse die Schlußfolgerung nahelegen, daß nur der Zinsspread die Zwei-Jahres Inflationsrate bedeutsam verändert, sollen die anderen beiden Kointegrationsbeziehungen im VECM berücksichtigt werden.

4.3.6 Out-of-Sample Prognosen anhand der konstruierten Vektorfehlerkorrekturmodelle

Die In-Sample Daten im Abschnitt zuvor umfassen den Zeitraum Januar 1973 bis Dezember 1996 zur Schätzung der VECM Modelle. Im Folgenden sollen diese Modelle zur Inflationsprognose anhand des Out-of-Sample Datensatzes, der den Zeitraum von Januar 1997 bis Dezember 2002 einschließt, angewendet werden. Ziel ist es, die Inflationsrate π_t^m in $m = 24$ Monaten vorherzusagen. Die Prognosen erfolgen nicht rekursiv, da sie als Referenzschätzungen im Zusammenhang mit den genetisch entwickelten Modellen dienen. Das vorliegende Kapitel untersucht primär verschiedene Strategien zur ex-anten Auswahl von Prognosegleichungen. Ein rekursives Schätzen der genetischen Modelle wäre ein extrem zei-

²³Auch der vom Statistischen Bundesamt zurückgezogene Konsumentenpreisindex könnte einen weiteren Grund darstellen.

Kointegrationsvektor	$\Delta\pi^0$	$\Delta\pi^2$	Δi^0	Δi^m
$i^0 - \pi^0 - c_{11}$	-0.110 (-4.5)	-0.022 (-1.3)	-0.046 (-3.0)	-0.070 (-4.0)
$i^1 - \pi^2 - c_{22}$	0.161 (5.5)	0.012 (0.6)	0.027 (1.5)	0.030 (1.4)
$i^1 - i^0 - c_{21}$	0.144 (3.1)	0.062 (1.9)	0.067 (2.3)	-0.017 (-0.5)
$i^0 - \pi^0 - c_{11}$	-0.104 (-4.2)	-0.020 (-1.1)	-0.039 (-2.5)	-0.061 (-4.0)
$i^2 - \pi^2 - c_{22}$	0.161 (5.4)	0.015 (0.7)	0.023 (1.3)	0.018 (1.0)
$i^2 - i^0 - c_{21}$	-0.139 (-3.5)	-0.056 (-2.0)	-0.061 (-2.5)	0.007 (0.3)
$i^0 - \pi^0 - c_{11}$	-0.104 (-4.1)	-0.02 (-1.1)	-0.039 (-2.5)	-0.054 (-3.8)
$i^3 - \pi^2 - c_{22}$	0.164 (5.3)	0.016 (0.8)	0.024 (1.2)	0.011 (0.7)
$i^3 - i^0 - c_{21}$	-0.136 (-3.7)	-0.05 (-2.0)	-0.057 (-2.5)	0.004 (0.2)
$i^0 - \pi^0 - c_{11}$	-0.104 (-4.1)	-0.019 (-1.1)	-0.041 (-2.6)	-0.05 (-3.8)
$i^4 - \pi^2 - c_{22}$	0.165 (5.3)	0.017 (0.8)	0.023 (1.2)	0.008 (0.5)
$i^4 - i^0 - c_{21}$	-0.134 (-3.9)	-0.048 (-2.0)	-0.053 (-2.4)	0.004 (0.2)
$i^0 - \pi^0 - c_{11}$	-0.105 (-4.1)	-0.018 (-1.0)	-0.042 (-2.7)	-0.047 (-3.8)
$i^5 - \pi^2 - c_{22}$	0.166 (5.3)	0.017 (0.8)	0.021 (1.1)	0.006 (0.4)
$i^5 - i^0 - c_{21}$	-0.133 (-3.9)	-0.046 (-2.0)	-0.049 (-2.4)	0.003 (0.2)

Tabelle 4.15: Anpassungsparameter der Kointegrationsvektoren; die t-Statistiken sind in Klammern angegeben.

tintensives Unterfangen. Um vernünftige rekursive Schätzungen anhand genetischer Programme zu tätigen, müßten für jeden zu prognostizierenden Monat eine Vielzahl an Läufen erzeugt werden. Betrachtet man die Prognosegleichungen der Vektorfehlerkorrekturmodelle und speziell die Gleichung der Änderungsrate der durchschnittlichen Inflationsrate, so erkennt man, daß die Gleichung unter anderem die beiden zeitverzögerten Terme sowohl der durchschnittlichen Inflationsrate der Differenzen als auch der Niveaus enthält. Da diese Daten nicht als realisierte beobachtbare Daten vorliegen, müssen sie prognostiziert werden. Als Veranschaulichung dient eine Gleichung aus dem System bei der ein Zinsspread zwischen einer Null-Kupon Anleihe mit einer Restlaufzeit von zwei Jahren und einem dreimonatigem Geldmarktzins die Differenz zwischen der durchschnittlichen zukünftigen zweijährigen Inflationsrate und der aktuellen vorhersagen soll:

$$\begin{aligned} \Delta\pi_t^{24} = & 0.02\pi_{t-1}^{24} - 0.015\pi_{t-1}^0 + 0.036i_{t-1}^2 - 0.041i_{t-1}^0 - 0.006 \\ & - 0.393\Delta\pi_{t-1}^{24} + 0.017\Delta\pi_{t-1}^0 - 0.036\Delta i_{t-1}^2 + 0.032\Delta i_{t-1}^0 \end{aligned} \quad (4.39)$$

Möchte man nun ab Januar 1997 Out-of-Sample Prognosen berechnen, so bedeutet dies Folgendes. Für die Prognose der zukünftigen Änderungsrate der durchschnittlichen Inflationsrate in 24 Monaten zum Zeitpunkt t $\Delta\pi_t^{24}$ liegen die Daten der Zinszeitreihen, die der Inflationsrate und der durchschnittlichen Inflationsrate bis zum Zeitpunkt t vor. Die VECM Gleichung für die Änderung der durchschnittlichen Inflationsrate enthält unter anderem folgende Terme²⁴:

$$\Delta\pi_t^{t+24} = \dots\Delta\pi_{t-1}^{t+23} \dots + \pi_{t-1}^{t+23} \quad (4.40)$$

²⁴Vgl. auch Kretschmer[76].

Modell	In-Sample 1975:3-1996:12	Out-of-Sample 1997:1-2002:12
i^1, π^2	0.708	1.740
i^2, π^2	0.706	1.603
i^3, π^2	0.708	1.961
i^4, π^2	0.710	2.159
i^5, π^2	0.711	2.334
<i>R.W.</i>	4.064	1.203

Tabelle 4.16: Fehlerabweichungen (MSE) bei Modellen mit Variation der Zinsen und Verwenden der durchschnittlichen 2-Jahres-Inflationsrate π^2 für deutsche Datensätze.

Da die beiden Terme auf der rechten Seite nicht als Beobachtungswerte vorliegen, müssen sie geschätzt werden. Erst für den Zeitpunkt $t - 23$ existieren realisierte Daten für diese beiden Terme, so daß $\Delta\pi_{t-23}^{t+1}$ berechnet werden kann:

$$\Delta\pi_{t-23}^{t+1} = \dots\Delta\pi_{t-24}^t \dots + \pi_{t-24}^t \quad (4.41)$$

Dieser Prognosewert wird in der nächsten Periode auf der rechten Seite eingesetzt, so daß man den nächsten Prognosewert $\Delta\pi_{t-22}^{t+2}$ erhält:

$$\Delta\pi_{t-22}^{t+2} = \dots\Delta\pi_{t-23}^{t+1} \dots + \pi_{t-23}^{t+1} \quad (4.42)$$

Durch das rekursive Schätzen gelangt man schließlich zum gewünschten Prognosewert für $\Delta\pi_t^{t+24}$. Die Berechnung der Schätzwerte für das Niveau der durchschnittlichen Inflationsrate erfolgt analog. Die Änderungsgraten müssen lediglich in Niveauvariablen umgerechnet werden. Hierfür soll die Gleichung zur Berechnung der durchschnittlichen Inflationsrate herangezogen werden. Da stets die zukünftige Inflationsrate in zwei Jahren geschätzt werden soll, dient als Ausgangspunkt folgende Gleichung:

$$\pi_t^{t+24} = \left(\sqrt{\left(1 + \frac{\pi_{t+24}}{100}\right)\left(1 + \frac{\pi_{t+12}}{100}\right)} - 1 \right) * 100 \quad (4.43)$$

Durch einfache mathematische Operationen und Auflösen der Gleichung nach π_{t+24} erhält man die geschätzte Inflationsrate in zwei Jahren:

$$\pi_{t+24} = \left(1 + \frac{\pi_t^{t+24}}{100}\right)^2 * \frac{100}{1 + \frac{\pi_{t+12}}{100}} - 100 \quad (4.44)$$

Man erkennt, daß neben dem Schätzwert der durchschnittlichen Inflationsrate auch der rekursiv ermittelte Schätzwert für die Ein-Jahres-Inflationsrate benötigt wird. Die Berechnung für π_{t+12} erfolgt analog:

$$\pi_{t+12} = \left(1 + \frac{\pi_{t-12}^{t+12}}{100}\right)^2 * \frac{100}{1 + \frac{\pi_t}{100}} - 100 \quad (4.45)$$

Modell	In-Sample 1975:3-1996:12	Out-of-Sample 1997:1-2002:12
Random Walk	4.064	1.203
ARIMA(2,1,0)	4.580	1.342
ARFIMA(1,d,1)		1.294
VECM (i^2, π^2)	0.706	1.603

Tabelle 4.17: Ergebnisse der Vergleichsmodelle

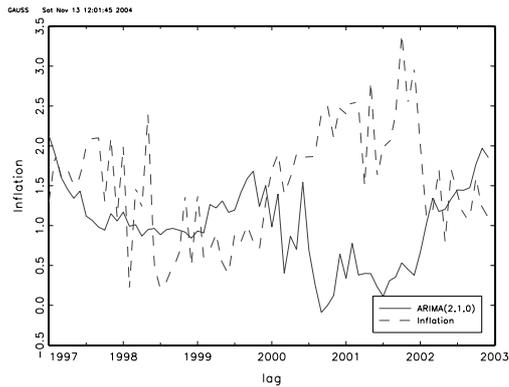
Anhand von π_{t+24} wird die mittlere quadratische Fehlerabweichung MSE für die Prognosegüte ermittelt. Die Tabelle 4.16 stellt die Ergebnisse der Vektorfehlerkorrekturmodelle bei Variation des Zinsspreads und Betrachtung der durchschnittlichen Zwei-Jahres-Inflationsrate dar. Während die In-Sample Schätzungen der VECM Modelle deutlich besser ausfallen als die Werte eines Random Walk Modells, ist keines der VECM Modelle in der Lage, den Random Walk bei der Out-of-Sample Prognose zu schlagen. Zudem zeigen die Ergebnisse, daß bei Zunahme der Restlaufzeiten der Null-Kupon Anleihen die Prognosegüte abnimmt.

4.4 Genetisch entwickelte Prognosemodelle

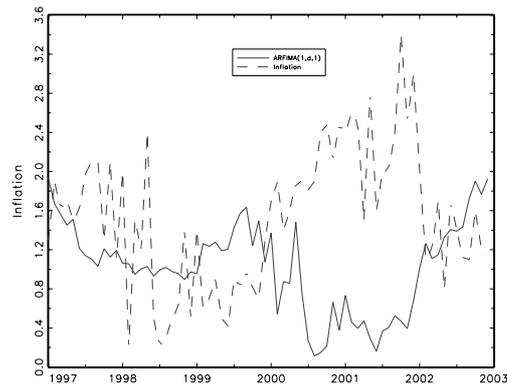
Nachdem in den vorangegangenen Abschnitten die In-Sample und Out-of-Sample Werte ökonomischer Vergleichsmodelle (Tabelle 4.17 und Abbildung 4.1) berechnet wurden, beschäftigt sich dieser Abschnitt mit der Untersuchung genetisch entwickelter Prognosemodelle. Das primäre Ziel dieses Kapitels besteht in der Überprüfung, ob bestimmte Kriterien bei der ex-ante Auswahl von Prognosegleichungen existieren, die die Prognosequalität erhöhen. Da die qualitativen Aussagen der Untersuchungen unabhängig vom gewählten Datensatz sind, konzentrieren sich die Anwendungen auf die Zeitreihen π^2, π^0, i^2, i^0 und somit auf die Kombination mit den besten Vorhersagewerten der Kointegrationsmodelle. Ein weiterer Grund für die Einschränkung der Datenwahl beruht auf dem nicht unerheblichen Zeitaufwand, der für die Entwicklung genetischer Modelle benötigt wird. Die Grundeinstellung der Parameterwerte des Genetischen Programms richtet sich weitestgehend nach den im Abschnitt zur Sensitivitätsanalyse erhaltenen Resultate und sind in Tabelle 4.18 aufgelistet. Bei den genetischen Operationen entfallen 60% auf Crossover- und 40% auf Mutationsoperationen. Zusätzlich wird der modifizierte Mutationstyp B verwendet, der Operanden durch neu erzeugte mathematische Teilstränge ersetzt. Während die Funktionsmenge die elementaren und trigonometrischen Operatoren enthält, berücksichtigt die Terminalmenge neben natürlichen und reellen Zahlen die Finanzmarktzeitreihen des zu untersuchenden Datensatzes, wobei das Programm Zeitverzögerungen bis zu 12 Monaten zuläßt:

$$\text{Funktionsmenge} = \{+ \ - \ / \ * \ \sin \ \cos \ \ln \ \exp \ \sqrt{n} \}$$

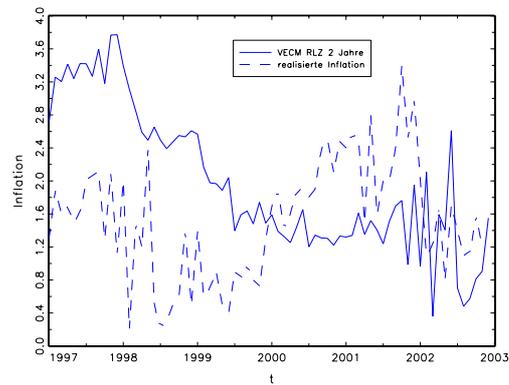
$$\text{Terminalmenge} = \{\text{Finanzmarktzeitreihen}, \mathbb{R} \in [0, 1], \mathbb{N} \in [1, 100]\}$$



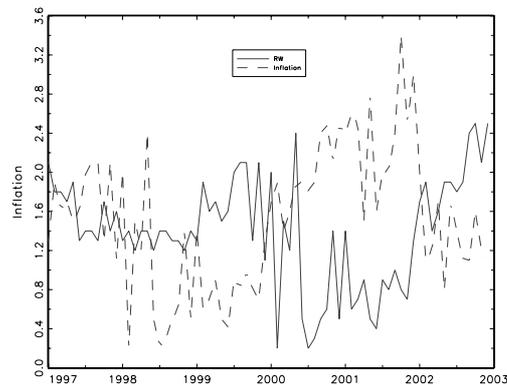
(a) Inflationsprognose - ARIMA(2,d,0)



(b) Inflationsprognose - ARFIMA(1,d,1)



(c) Inflationsprognose - VECM RLZ 2 Jahre



(d) Inflationsprognose - Random Walk

Abbildung 4.1: Out-of-Sample Prognosen der Vergleichsmodelle.

Zugrundegelegt wird eine proportionale Selektionsmethode, deren Auswahlwahrscheinlichkeit vom Fitneßwert der Lösungsgleichungen bestimmt wird.

Die Überprüfung verschiedener Methoden zur ex-ante Auswahl der Prognosegleichungen orientiert sich am folgenden Untersuchungsschema:

Die Untersuchungen beginnen zunächst mit einer simplen Aufteilung der Daten in eine In-Sample und eine Out-of-Sample Menge, wobei die Auswahl der Prognosegleichungen ausschließlich über die In-Sample Fehlerabweichung erfolgt. Anschließend wird eine Dreiteilung der Datenmenge vorgenommen, die durch das Zwischenschalten einer Trainingsphase gekennzeichnet ist. Mit Hilfe der Trainingsphase kann die Prognosegüte der entwickelten Gleichungen überprüft werden und als Selektionskriterium der Gleichungen für die eigentliche Prognose anhand der Out-of-Sample Daten dienen. Diese Methode stellt ein in der Literatur übliches Verfahren dar. Des weiteren werden die in der Sensitivitätsanalyse entwickelten Module des Backpropagation Verfahrens sowie das Analyseverfahren der Building Blocks unter Berücksichtigung einer integrierten Trainingsphase angewendet. Es soll überprüft werden, wie sich das zusätzliche Einschalten des Optimierungsverfahrens

Anzahl der Läufe	=	100
Anzahl der Generationen pro Lauf	=	1000
Populationsgröße	=	500
Fitneßkriterium	=	MSE (Variante 5: AIC)
maximale Länge der Funktionen der Anfangsgpopulation	=	5 Knoten
maximale Länge der Funktionen im weiteren Verlauf	=	100 Knoten
Anzahl neuer Gleichungen durch Crossoveroperationen	=	300
Anzahl neuer Gleichungen durch Mutationsoperationen	=	200
Typ A:	=	100
Typ B:	=	100
maximale Länge der Funktionsteile bei Mutationstyp B	=	5 Knoten
Anzahl der besten Funktionen, die für genetische Operationen herangezogen werden	=	100

Tabelle 4.18: Anfangsparameter des GP GAUSS Grundmodells
Mutationstyp A: Operand wird durch einen Operand und Operator durch einen Operator ersetzt. Mutationstyp B: Austausch eines Operanden durch einen neu generierten Funktionsteil.

rens sowohl für die In-Sample als auch Out-of-Sample Schätzung auswirkt. Die Intention, die hinter der Anwendung der Building Block Analyse steht, zielt in eine etwas andere Richtung. Das entwickelte Verfahren der Building Block Analyse bildet die Aktivität des Genetischen Programms ab, indem es nicht nur die Anzahl erfolgreicher genetischer Operationen sondern auch durch das Anzeigen der Fitneßänderungen deren Ausmaß offenlegt. Der Übergang von einer aktiven zu einer ruhigen Phase soll als Auswahlkriterium dienen. Im letzten Abschnitt des Kapitels wird das Fitneßkriterium der mittleren quadratischen Fehlerabweichung (MSE) durch ein Kriterium ähnlich dem Akaike Informationskriterium ersetzt, so daß das Vorgehen insgesamt analog zur Sensitivitätsanalyse erfolgt. Beim Informationskriterium geht die Anzahl der involvierten Regressoren negativ in die Bewertung der Prognosegleichungen ein. Neben einer Punktprognose anhand der mittleren quadratischen Fehlerabweichung (MSE) werden zusätzlich Richtungsprognosen erstellt, deren Anforderungen geringer ausfallen als beim MSE, da die Prognosemodelle im Gegensatz zum MSE nur die Richtung nicht aber das Ausmaß korrekt vorhersagen müssen.

4.4.1 In-Sample und Out-of-Sample

Die erste Modellvariante beginnt mit einer Zweiteilung der Daten in eine In-Sample und eine Out-of-Sample Periode, die sich über einen Zeitraum von 1975:3-1996:12 und 1997:1-2002:12 erstrecken. Anhand der In-Sample Daten schätzt das Genetische Programm die Prognosegleichung für die zukünftige Inflationsrate in zwei Jahren. Insgesamt generiert GP GAUSS aufgrund der festgelegten Parameter 50 Millionen potentielle Prognosegleichungen²⁵, wobei am Ende jeder Generation nur die beste Gleichung gespeichert wird.

²⁵500 Gleichungen pro Generation \times 1000 Generationen \times 100 Läufe = 50 Mio. Gleichungen

Auf der Grundlage von insgesamt 100.000 Lösungskandidaten werden verschiedene Auswahlstrategien angewendet und miteinander verglichen. Als erstes werden die Gleichungen betrachtet, die die zu regressierende Inflationszeitreihe am besten approximieren. Dazu soll aus den einhundert Läufen derjenige mit der kleinsten quadratischen Fehlerabweichung und aus diesem Lauf wiederum die Funktionen nach 200 und 1000 Generationen für die Prognose gewählt werden. Die in Tabelle 4.19 aufgeführten Ergebnisse zeigen, daß

Modell	Generation	In-Sample	Out-of-Sample	Knoten	Lauf
beste Funktion	200	0.455	2.617	94	4
besten 10 Funktionen	200	0.532	13.239	94.8	
alle Funktionen	200	0.712	641.887	93.5	
beste Funktion	1000	0.440	1.741	99	66
besten 10 Funktionen	1000	0.478	38.083	98.8	
alle Funktionen	1000	0.598	4240581.3	97.2	

Tabelle 4.19: In-Sample, Out-of-Sample; die Ergebnisse wurden ausgehend von den besten In-Sample Regressionsgleichungen ermittelt.

die In-Sample Schätzungen nach 1000 Generationen sowohl für die beste Funktion als auch für den Durchschnitt der besten zehn Funktionen deutlich besser ausfallen als die Ergebnisse der Benchmarkmodelle. Gegenüber dem VECM Modell mit einem MSE Wert von 0.706 ergibt sich eine Verringerung der Fehlerabweichung von 37.7%. Der Blick auf die Out-of-Sample Vorhersagen sorgt allerdings für Ernüchterung. Keines der hier berechneten Ergebnisse schneidet im Vergleich zu den Benchmarkmodellen besser ab. Abgesehen von den bereits schwachen Ergebnissen der besten Funktion, verschlechtern sich die Prognosen durch Bilden von Durchschnittswerten erheblich. Der Vergleich zwischen der 1000. und 200. Generation mit der nochmaligen Verschlechterung der Durchschnittswerte kann als ein Indiz für Overfitting interpretiert werden. Im Gegenzug fallen die In-Sample Verbesserungen der besten Regressionsgleichung mit 3.3% nur marginal aus, so daß eine Ausweitung der Generationenanzahl auf 1000 wenig sinnvoll erscheint. Unterstützt wird

Ergebnis	MSE nach		prozentuale Verbesserung
	200 Generationen	1000 Generationen	
1	2.040	0.636	68.8
2	1.485	0.486	67.3
3	1.278	0.530	58.5
4	1.371	0.640	53.3
5	1.201	0.667	44.4
6	0.455	0.440	3.4
7	0.709	0.598	11.9

Tabelle 4.20: MSE Entwicklung der größten prozentualen Verbesserungen aus Abbildung 4.2.

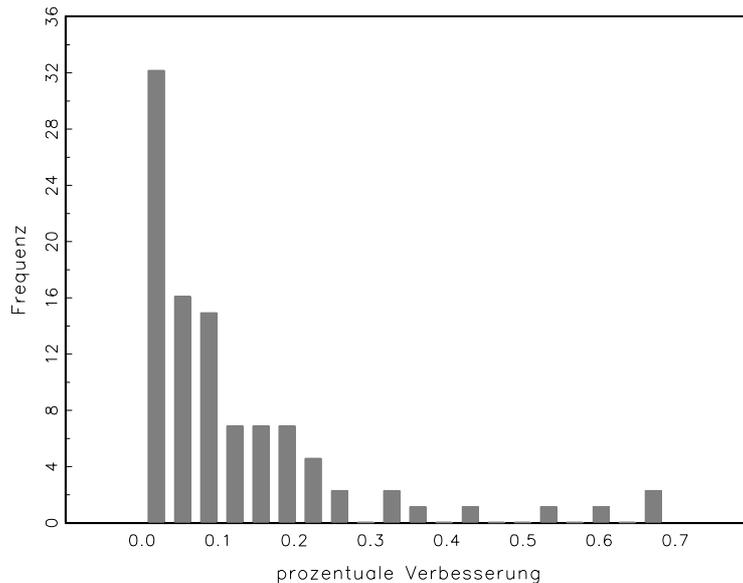


Abbildung 4.2: Prozentuale Verbesserung zwischen Generation 200 und 1000.

diese Vermutung durch die Abbildungen 4.2 bis 4.4 sowie der Tabelle 4.20. Abbildung 4.2 zeigt, daß bei den meisten Läufen bereits nach 200 Generationen keine wesentlichen Verbesserungen mehr eintreten. Bei über 50% der Läufe betragen die Verbesserungen weniger als 10%. Die Verbesserungen der übrigen Läufe von teilweise 70% lassen sich mit der Tatsache begründen, daß die Fitneßwerte nach 200 Generationen im Vergleich zum Durchschnitt erheblich schlechter ausfallen, wie Tabelle 4.20 veranschaulicht (Ergebnisse 1 bis 5). Ungeachtet der geringen Verbesserung der besten Regressionsgleichung von 3.3% und der gleichzeitigen Verschlechterung der durchschnittlichen Prognosewerte (Tabelle 4.19) wird die durchschnittliche In-Sample Verbesserung von 16.0%, die man bei einer Verfünffachung der Generationenanzahl erzielt, durch einen erheblichen zusätzlichen Zeitaufwand erkauft.

Ein letzter Beleg für die kontraproduktive Ausweitung der Generationenanzahl liefern die beiden Abbildungen 4.3 und 4.4. Grundlage der beiden Graphiken bilden die Prognosewerte der jeweils 100.000 generierten Schätzgleichungen, wobei gleichwertige Prognosen eliminiert und MSE-Prognosewerte größer als 2.0 durch diesen Wert für eine bessere Veranschaulichung substituiert wurden. Bei einer Genauigkeit von neun Nachkommastellen kann davon ausgegangen werden, daß ausschließlich die Prognosen identischer Schätzgleichungen gelöscht wurden. Die Abbildungen zeigen, daß innerhalb der ersten 200 Generationen nur ungefähr 38% schlechtere Ergebnisse als einen MSE-Wert von 2.0 aufweisen im Vergleich zu ungefähr 45% zwischen Generation 200 und 1000. Auch die Anzahl besserer Prognosewerte im Vergleich zu einem Random Walk (MSE-Wert = 1.203) ist innerhalb der ersten 200 Generationen deutlich größer. Insgesamt betrachtet können die letzten beiden Aspekte als weitere Hinweise für das Overfitting Problem aufgefaßt werden. Abgesehen von den schlechten Out-of-Sample Ergebnissen in Tabelle 4.19 und den Indizien für das Overfitting Problem dokumentieren die Abbildungen 4.3 und 4.4 allerdings auch, daß genetisch entwickelte Modelle überhaupt in der Lage sind, bessere Vorhersagen als ein naives

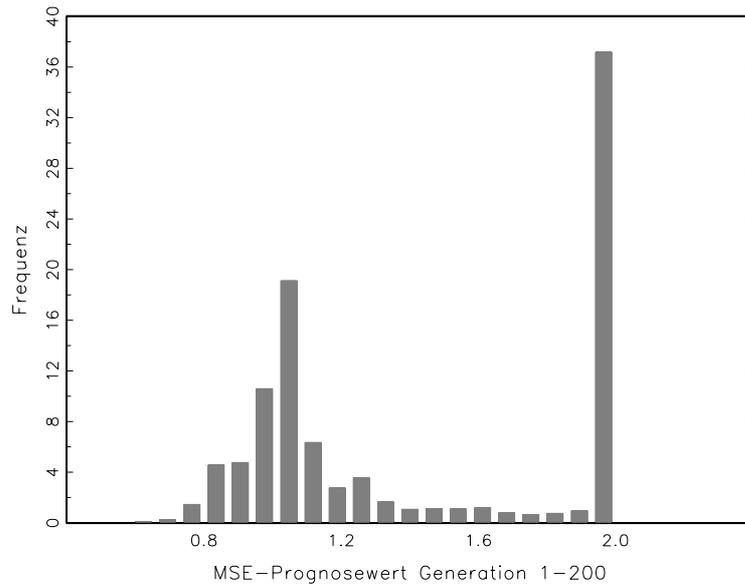


Abbildung 4.3: MSE Prognosewerte zwischen Generation 1 und 200.

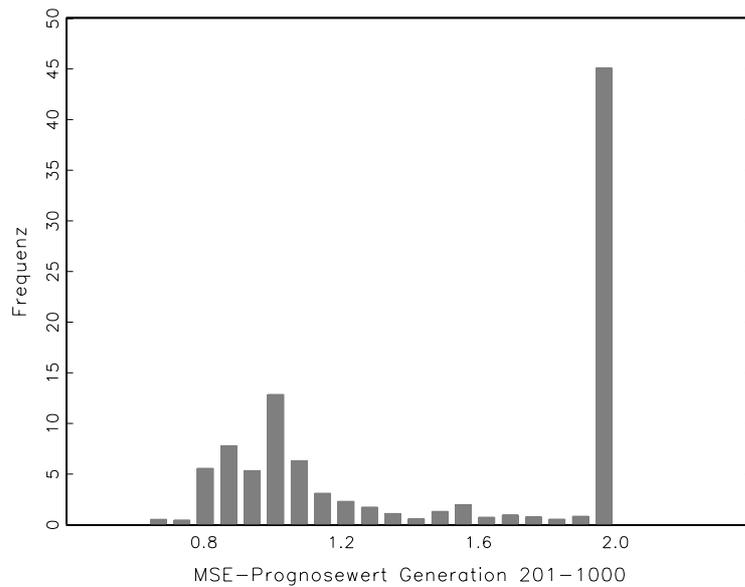


Abbildung 4.4: MSE Prognosewerte zwischen Generation 1 und 1000.

Modell zu erzeugen. Tabelle 4.21 listet die Werte für die beiden Regressionsgleichungen mit den besten Out-of-Sample Resultaten auf, die in einer ex-post Betrachtung bei dieser Simulation maximal erreicht werden konnten²⁶. Mit Werten von 0.588 für die ersten 200 Generationen und 0.673 für die restlichen Generationen existieren Prognosewerte, die weit unterhalb eines Random Walks liegen. Der modifizierte Diebold-Mariano Test zeigt jedoch, daß die Out-of-Sample Prognosen nicht signifikant vom Random Walk abweichen.

²⁶Die Out-of-Sample Prognosen sind in Abbildungen 4.7 und 4.8 dargestellt.

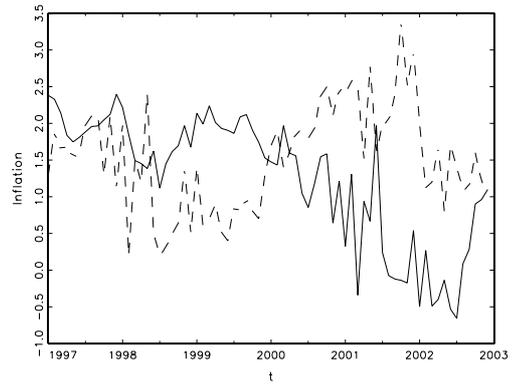
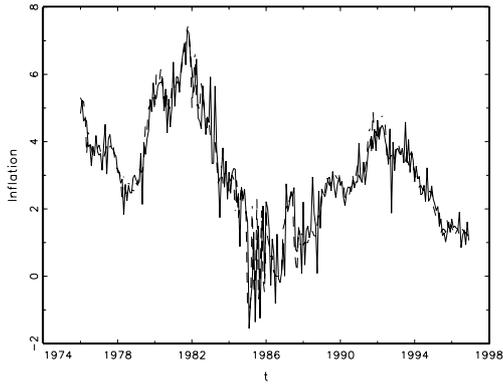


Abbildung 4.5: Tabelle 4.19: Lauf 66, Generation 1000

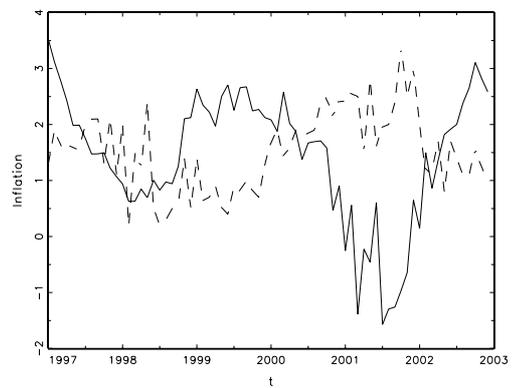
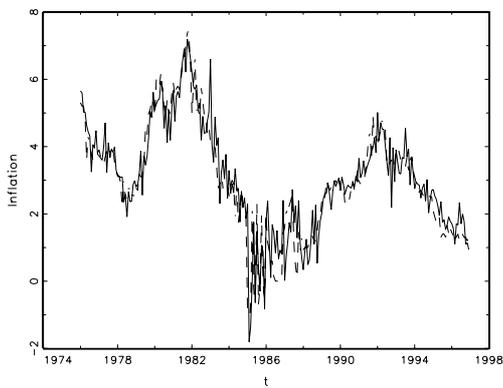


Abbildung 4.6: Tabelle 4.19: Lauf 4, Generation 200

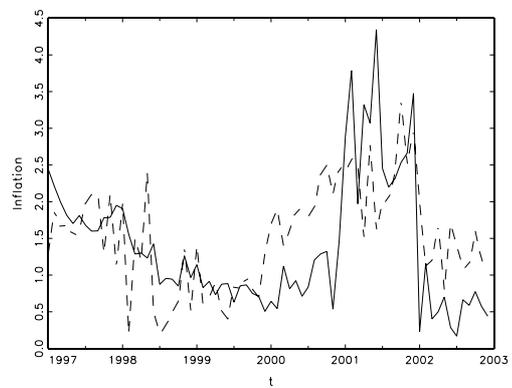
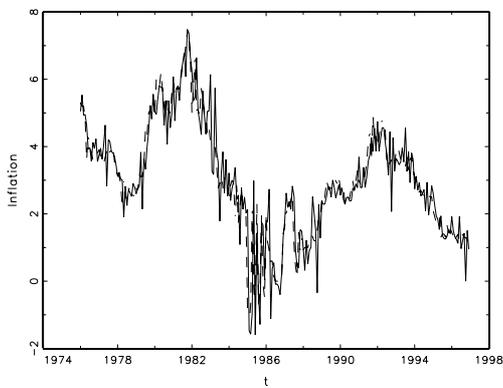


Abbildung 4.7: Tabelle 4.21: Lauf 21, Generation 1000 (Modell B) - ex-post beste Funktion

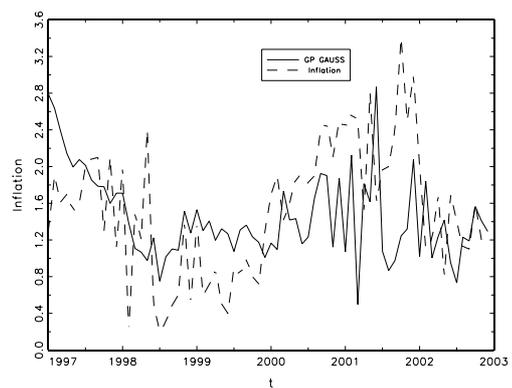
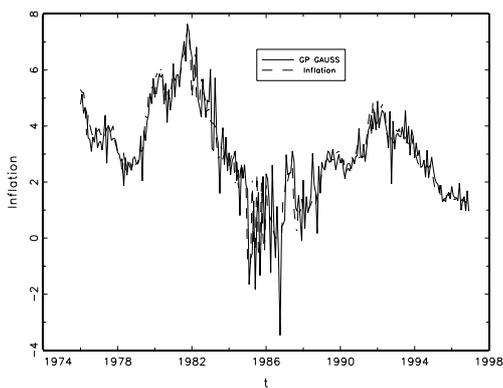


Abbildung 4.8: Tabelle 4.21: Lauf 66, Generation 200 (Modell A) - ex-post beste Funktion

Modell	Generation	Out-of-Sample	In-Sample	Knoten	Lauf
A: beste Funktion	200	0.588 (0.74)	0.567	99	66
B: beste Funktion	1000	0.673 (0.77)	0.491	99	21

Tabelle 4.21: In-Sample und Out-of-Sample; die besten Out-of-Sample Ergebnisse. Die Signifikanzniveaus des Diebold-Mariano Tests sind in Klammern angegeben. Demnach weichen die Prognosewerte nicht signifikant vom Random Walk Prognosewert ab.

Die folgende Funktion gibt die Prognosegleichung für Lauf 66 und Generation 200 wieder:

$$\begin{aligned}
\pi_t^{24} = & (\hat{\pi}^{24}[t-1] - (\sin(((\sin(t) + ((\sin(40) + ((\sin(\pi^0[t-6] \\
&) + ((\sin(t) + ((\sin(40) + ((\sin(\pi^0[t-6]) + ((\sin(i^2[t-9] \\
&) + ((0.69 - \text{abs}(i^0[t-4])/0.75))/\hat{\pi}^{24}[t-2] + (\pi^0[t-10] - \\
& \text{exp}(\pi^0[t-10]/i^0[t-4])) - \hat{\pi}^{24}[t-1] + \pi^0[t-10]))) \\
&)))))/ (40 - (\text{abs}(i^0[t-4]) / (\text{wzl}((\text{abs}(((\pi^0[t-10] - (\text{abs}(\\
& (i^0[t-4] / (97 / (\pi^0[t-10] / i^0[t-4]) * i^0[t-4] \\
& * 90)))))) - i^0[t-4])))
\end{aligned} \tag{4.46}$$

Gleichung 4.46 stellt nicht die original ausgeworfene Prognosegleichung von GP GAUSS dar. Es wurde versucht, die überflüssigen Elemente zu beseitigen. Im Anhang auf der Seite 154 befindet sich die ursprüngliche Prognosegleichung zum Vergleich. Die Schätzung des Terms $\hat{\pi}_t^{24}$ entspricht dem Vorgang aus Abschnitt 4.3.6. Demnach enthält die genetisch entwickelte Prognosegleichung alle erklärenden Variablen mit bis zu zehn Zeitverzögerungen. Ein letzter Aspekt im Zusammenhang mit der Overfitting Problematik wird in Abbildung 4.9 graphisch dargestellt. Die in der Graphik abgebildeten MSE Werte eines ausgewählten Laufs wurden durch Division der Random Walk Werte normiert, so daß Werte größer als eins schlechtere Prognosen als ein Random Walk ausdrücken. Prognosewerte größer als MSE = 2 wurden erneut auf den Wert zwei gestutzt, um die Graphik anschaulicher zu gestalten. Darüber hinaus wurde die Entwicklung der Knotenanzahl ebenfalls skaliert. In der Literatur wird das Problem des Overfitting als eine allgemeine Erscheinung behandelt. Es soll an dieser Stelle betont werden, daß das Overfitting Problem nur im Durchschnitt auftaucht, aber auf der Ebene einzelner Gleichungen keine Allgemeingültigkeit besitzt. Einzelne Läufe, wie zum Beispiel Abbildung 4.9, zeigen im Zeitablauf immer wieder Schwankungen im Hinblick auf die Vorhersagegüte ohne sich tendenziell zu verschlechtern. Dagegen ist die Konvergenz der Knotenlänge, die die Länge der Regressionsgleichungen widerspiegelt, hin zur maximal möglichen Größe bei zugrundelegen des MSE Fitneßkriteriums sowohl im Durchschnitt als auch auf der Ebene einzelner Gleichungen ein allgemeingültiges Phänomen. Zum Schluß dieses Abschnittes richtet sich der Blick auf die Richtungsprognosen. Tabelle 4.22 enthält sowohl die Werte der einzelnen Vergleichsmodelle als auch der beiden besten ex-post Prognosegleichungen. Die Prognosegüte des Random Walks fällt im Hinblick auf die Richtungsänderung der Inflationsrate mit einer Trefferquote von 35% gering aus, so daß die Hürde für die anderen Modelle relativ niedrig ist. Obwohl das Genetische Programm in der ex-post Betrachtung bei der Punktprognose mit

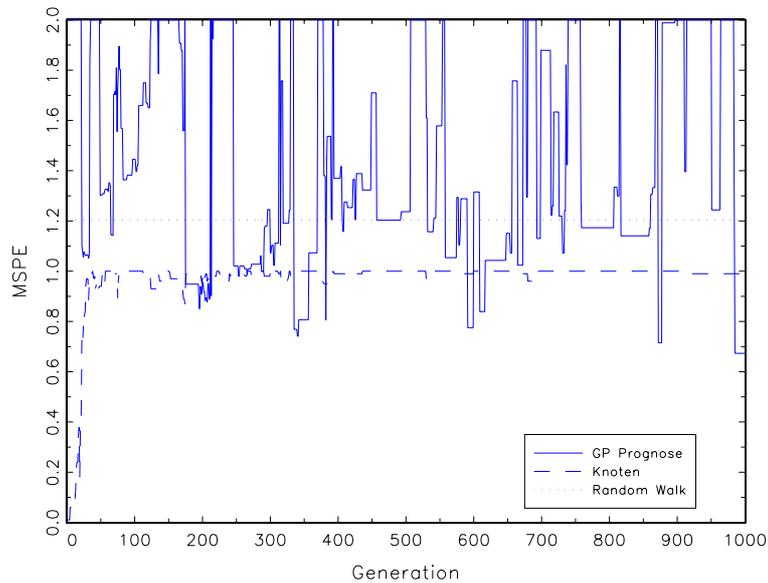


Abbildung 4.9: Prognosegüte der genetisch entwickelten Gleichung aus dem Lauf 21 im Zeitablauf. Skalierung der Knotenentwicklung: der Wert eins entspricht 100 Knoten.

Abstand die besten Werte erzielt, erreicht es abgesehen vom Random Walk nur mit Mühe die Prognosefähigkeit der anderen Vergleichsmodele. Trefferquoten von 56% und 50% stehen Werte von 60% beim ARIMA und ARFIMA Modell gegenüber. Allerdings sind alle Richtungsprognosen eindeutig nicht signifikant.

Insgesamt kann festgehalten werden, daß gute Punktprognosen mit dieser Auswahlstrategie nur zufällig ausgewählt werden. Darüber hinaus sind die analysierten Prognosewerte nicht signifikant. Auch die Richtungsprognosen sind nur bedingt positiv zu beurteilen. Zwar schneiden sie im Vergleich zum Random Walk wesentlich besser ab, nicht jedoch gegenüber den anderen Vergleichsmodellen. Zudem besteht das Problem, daß die Richtungsprognosen ebenfalls nicht signifikant sind.

Modell	Trefferquote in %	Teststatistik	Signifikanz
A	56.33	0.643	0.42
B	50.70	0.006	0.94
ARIMA	60.56	1.566	0.21
ARFIMA	60.56	1.596	0.21
VECM	54.93	0.368	0.54
R.W.	35.21	3.097	0.08

Tabelle 4.22: Richtungsprognosen. Die anhand eines χ^2 -Tests ermittelten Signifikanzniveaus zeigen, daß die Trefferquoten nicht systematisch von einer Zufallsprognose abweichen.

4.4.2 In-Sample, Trainingsphase und Out-of-Sample

Nachdem die bisherigen Untersuchungen enttäuschende Prognoseergebnisse hervorgebracht haben, soll das Programm um eine zweijährige Trainingsphase, die sich über den Zeitraum von Januar 1995 bis Dezember 1996 erstreckt, erweitert werden. Dementsprechend verkürzt sich die In-Sample Periode auf März 1975 bis Dezember 1994 bei einer gleichgebliebenen Out-of-Sample Periode von Januar 1997 bis Dezember 2002. Der Einbau einer Trainingsphase ermöglicht dem Anwender, die anhand der In-Sample Daten genetisch entwickelten Lösungsgleichungen in einer ex-ante Prognose zu testen, um anschließend die beste Regressionsgleichung auf die Out-of-Sample Periode anzuwenden. Bei der Auswahl der Gleichungen werden im Folgenden vier verschiedene Strategien untersucht. Aus dem Lauf mit dem kleinsten In-Sample Fitneßwert wird die Funktion mit dem besten Prognosewert in der Trainingsphase ausgewählt. Zusätzlich werden aus diesem Lauf die Durchschnittswerte für die besten zehn Funktionen berechnet. Die Ergebnisse in Tabelle 4.23 ähneln qualitativ der ersten Modellvariante und weisen mit Werten von 8.385 und 14.335 keine Prognosefähigkeit auf. Vollständigkeitshalber ist erneut der Prognosewert der besten In-Sample Funktion aufgeführt und signalisiert mit einem MSE-Wert von 32.173 in einem noch stärkerem Ausmaß die Ungeeignetheit dieser Strategie. Eine weitere Möglichkeit, die sich aufgrund der Trainingsphase eröffnet, ist die Wahl der Prognosegleichung mit dem global besten Trainingswert sowie der Durchschnitt der besten zehn globalen Gleichungen. Beide Strategien schlagen mit Fehlerabweichungen von 1.007 und 0.976 den Random Walk. Allerdings belegt der Test von Diebold und Mariano erneut, daß die Abweichung vom Random Walk für die Gleichung mit dem globalen Minimum mit einem P-Wert von 0.55 nicht signifikant ist. Wie die Ergebnisse des nächsten Abschnitts zeigen werden, handelt es sich um zufallsbedingte Werte. Die in Tabelle 4.24 dargestellten Richtungsprognosen liefern keine neuen Erkenntnisse.

4.4.3 In-Sample, Trainingsphase, Out-of-Sample plus Backpropagation

Obwohl das Implementieren des Backpropagation Moduls auf der Suche nach einer optimalen Auswahlstrategie aufgrund der vorliegenden Ergebnisse sicherlich keine neuen Er-

Modell	Generation	In-Sample	Training Minimum	Out-of-Sample	Knoten	Lauf
beste Funktion ¹	1000	0.426	0.199	32.173	98	31
beste Funktion ²	258	0.536	0.100	8.385	98	31
besten 10 Funktionen ¹	1 bis 1000	0.551	0.103	14.335	93.2	
globales Minimum ³	302	0.597	0.094	1.007 (0.55)	99	29
globale 10 Minima ³	1 bis 1000	0.599	0.095	0.976 (0.53)	99.5	

Tabelle 4.23: ¹ aus In-Sample-Phase; ² aus Lauf 31 Funktion mit minimalem Training MSE; ³ aus Trainingsphase. Der in Klammern angegebene Wert ist das Signifikanzniveau des Diebold-Mariano Tests.

Modell	Lauf	Generation	Trefferquote	Teststatistik	Signifikanz
beste Funktion	31	1000	47.887	0.065	0.80
beste Funktion	31	258	47.887	0.105	0.75
besten 10 Funktionen	31	1 bis 1000	53.521	0.174	0.68
globales Minimum	29	302	52.113	0.058	0.81
globale 10 Minima		1 bis 1000	50.706	0.006	0.94

Tabelle 4.24: Richtungsprognose für Out-of-Sample Periode; Trefferquote in %

Modell	Generation	In-Sample	Training Minimum	Out-of-Sample	Knoten	Lauf
beste Funktion ¹	200	0.559	1.636	6935.159	100	47
beste Funktion ²	42	0.732	0.114	0.865 (0.63)	91	47
besten 10 Funktionen ¹	1 bis 200	0.757	0.127	1.044 (0.67)	50	47
globales Minimum ³	110	0.752	0.100	2.266	100	19
globale 10 Minima ³	1 bis 200	0.748	0.102	26.907	92.7	

Tabelle 4.25: Backpropagation; ¹ aus In-Sample-Phase; ² aus Lauf 47 Funktion mit minimalem Training MSE; ³ aus Trainingsphase. In Klammern: Signifikanzniveau des Diebold-Mariano Tests.

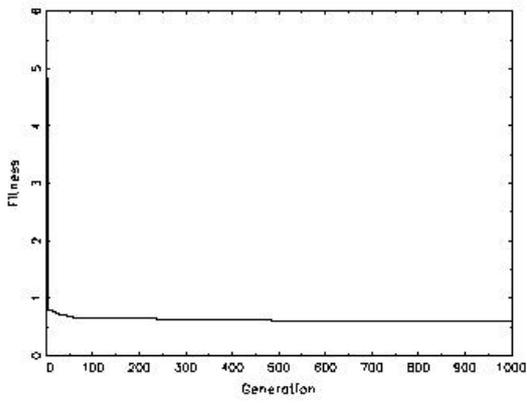
kenntnisse mit sich bringen wird, soll das Optimierungsmodul angewendet werden, um die Wirkung vor allem bei der In-Sample Schätzung zu überprüfen. Die strategische Auswahl der Prognosegleichungen wird analog zum vorherigen Abschnitt vorgenommen. Aufgrund der starken Rechenintensität wird die Anzahl der Generationen auf 200 begrenzt, wobei der auftretende Informationsverlust wegen der vorangegangenen Ergebnisse im Zusammenhang mit der Overfitting Problematik vernachlässigbar ist. Vergleicht man den besten In-Sample Wert von 0.559 mit dem entsprechenden Wert ohne Anwendung des Backpropagation Verfahrens 0.523 aus der Ergebnismatrix des vorherigen Abschnitts, dann muß man feststellen, daß das Optimierungsverfahren trotz erfolgreicher Anwendungen im Zusammenhang mit der Sensitivitätsanalyse (Kapitel 2.7) im Vergleich ohne Implementierung des Moduls zumindest in diesem Fall nicht in der Lage war, die In-Sample Ergebnisse weiter zu verbessern. Auffällig ist allerdings der sehr schlechte Prognosewert, der sich sogar im Vergleich zu anderen schlechten Prognosen deutlich abhebt. Ansonsten bleibt festzuhalten, daß sich die Ergebnisse in Tabelle 4.25 im Vergleich zu Tabelle 4.23 ins Gegenteil umkehren. Demnach erweist sich die Wahl der Gleichungen mit den global minimalen MSE-Werten aus der Trainingsphase als schlechte und die Wahl der besten Trainingsgleichungen aus dem Lauf mit dem minimalen In-Sample Fitneßwerten als gute Strategie. Die in Tabelle 4.26 dargestellten Richtungsprognosen liefern erneut keine neuen Erkenntnisse.

Modell	Lauf	Generation	Trefferquote	Teststatistik	Signifikanz
beste Funktion	47	200	57.746	1.017	0.31
beste Funktion	47	42	47.887	0.058	0.81
besten 10 Funktionen	47	1 bis 200	52.113	0.073	0.79
globales Minimum	19	110	50.704	0.012	0.91
globale 10 Minima		1 bis 200	51.850	0.020	0.89

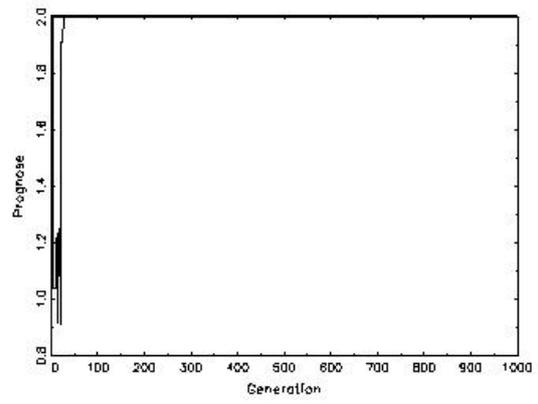
Tabelle 4.26: Backpropagation; Richtungsprognose für Out-of-Sample Periode; Trefferquote in %

4.4.4 In-Sample, Trainingsphase, Out-of-Sample und Building Blocks

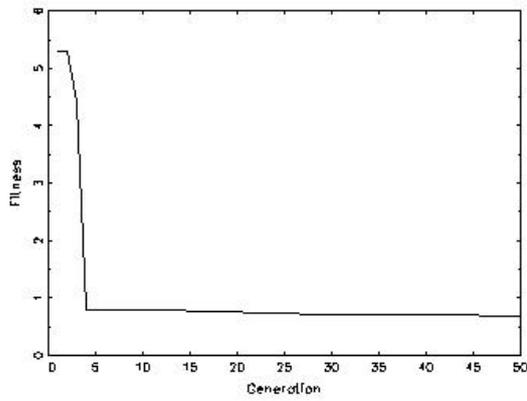
Die Auswahlstrategien der beiden folgenden Abschnitte setzen nicht primär bei den MSE Ergebnismatrizen der In-Sample- oder Trainingswerten an, sondern versuchen die *Entwicklung* genetischer Gleichungen im Zeitablauf zu berücksichtigen. Der Grund für das wechselhafte Abschneiden könnte in einem etwas anders ausgeprägten Overfitting Problem bestehen. Während es üblicherweise derart definiert ist, daß bei einer kontinuierlichen In-Sample Verbesserung die Prognosegüte sich kontinuierlich verschlechtert, könnte man es bei den vorliegenden Untersuchungen vielmehr als stufenweises Overfitting Problem bezeichnen. Spürt GP GAUSS einen nützlichen Building Block auf, so kommt es zu größeren Fitneßänderungen. Die anschließenden kleinen Fitneßänderungen könnte man als Feinjustierung der gefundenen Regressionsgleichung beschreiben. Diese Feinjustierung wiederum beinhaltet eine zunehmende Regression stochastischer Elemente. Um dies zu veranschaulichen, soll die im Abschnitt der Sensitivitätsanalyse bereits vorgestellte Untersuchungsmethode der Building Blocks zur Anwendung kommen. Die Building Block Analyse ermöglicht es, die Fitneßänderungen erfolgreicher genetischer Operationen darstellen zu können. Somit bekommt man eine graphische Analyse des Genetischen Programms in Bezug auf produktive Phasen mit großen Output-Änderungen und den Übergang zu ruhigeren Phasen mit marginalen oder keinen Fitneßänderungen. Die Abbildungen zeigen die Entwicklungen der In-Sample-, der Trainings- und der Out-of-Sample-Phase sowie die Entwicklung der Fitneßänderungen genetischer Operationen. Interessant sind die Phasen mit stärker ausgeprägten sprunghaften Verbesserungen der In-Sample Schätzungen. Sie signalisieren, daß GP GAUSS einen nützlichen Building Block gefunden hat, der einen neuen Prozeß mit Fitneßänderungen innerhalb der Population anstößt. Durch die Erneuerungen der Populationsmenge, die als Pool für genetische Operationen zur Verfügung steht, entstehen interdependente Änderungsprozesse, wodurch auch kleine Fitneßverbesserungen bei der besten Populationsgleichung die Prognose positiv beeinflussen könnten. Wie man aus den Abbildungen entnehmen kann, herrscht dieser Prozeß eine gewisse Zeit vor und ist den Entwicklungen der In-Sample und Trainingsphase zeitlich nachgelagert (Abbildungen 4.10). Es soll im Folgenden überprüft werden, ob der Zeitpunkt des Übergangs von der systematischen zur unsystematischen Approximation, also der Zeitreihe inhärenten stochastischen Struktur, besser abgeschätzt werden kann. Letztere wird häufig als Grund für das Overfitting Problem angesehen. Dazu wird die Generation überprüft,



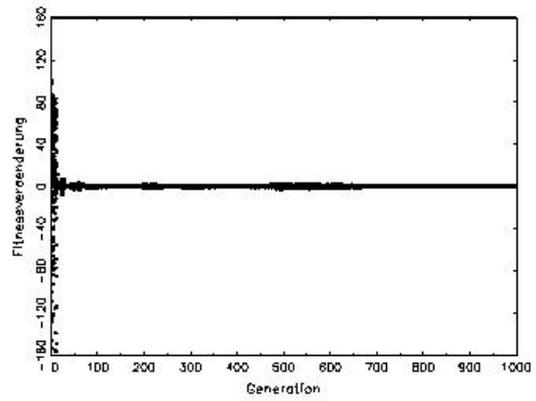
(a) In-Sample: Fitneßentwicklung



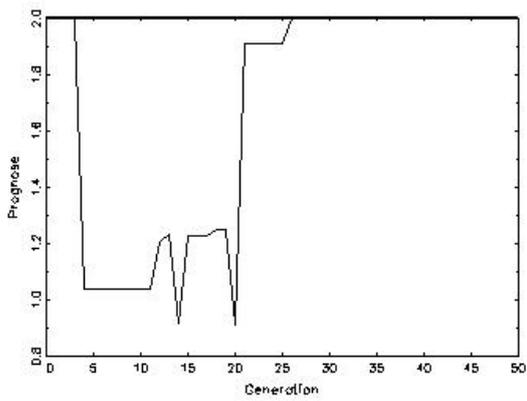
(b) Out-of-Sample: Prognoseentwicklung



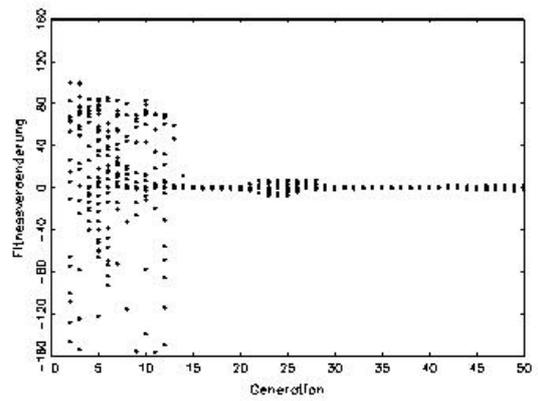
(c) In-Sample: Generation 1 bis 50



(d) Fitneßänderungen durch Crossover



(e) Out-of-Sample: Generation 1 bis 50



(f) Fitneßänderungen durch Crossoveroperationen: Generation 1 bis 50

Abbildung 4.10: Building Block Analyse: Phase der Fitneßänderungen durch Crossover ist der In-Sample Fitneßentwicklung zeitlich nachgelagert.

Gen	In-Sample	Training	Prognose	Strategie A ¹	Strategie B ²
3	4.415	5.069	2.0	99.816	24.599
4	0.787	0.136	1.039	84.673	18.862
10	0.787	0.136	1.039	69.872	21.643
11	0.786	0.135	1.040	0.434	6.520
12	0.784	0.138	1.205	0.458	9.177
13	0.780	0.129	1.230	46.641	1.165
14	0.777	0.130	0.917	0.996	0.243
15	0.774	0.128	1.225	1.483	0.133
16	0.774	0.128	1.225	0.193	0.115

Tabelle 4.27: ¹ absolute MSE Änderungen der besten In-Sample Gleichung; ² Durchschnitt absoluter MSE Änderungen sämtlicher Gleichungen.

bei der die Phase volatiler Fitneßänderungen aufhört und in eine ruhigere Phase eintritt. Konkret werden sowohl die absoluten Fitneßänderungen (Strategie A) der besten sowie der Durchschnitt aller einhundert Gleichungen (Strategie B) berechnet, um anschließend die Prognosegleichung derjenigen Generation zu wählen, bei der die absolute Änderungsrate die 1%-Marke unterschreitet, wobei dieser Grenzwert willkürlich getroffen wurde. In Tabelle 4.27 ist der Ausschnitt eines Laufs über mehrere ausgewählte Generationen abgebildet. Weder die In-Sample- noch die Trainingswerte geben Aufschluß über mögliche gute Prognosegleichungen. Da sich die In-Sample Werte kontinuierlich verbessern, bieten sich dem Anwender kaum Möglichkeiten unterschiedliche Strategien zu untersuchen. Die Entwicklung der Trainingswerte läßt ebenfalls keinen Zusammenhang mit den Prognosewerten erkennen. Die Fitneßänderungen der besten Prognosegleichung bilden die Fitneßverbesserung derjenigen Gleichung ab, die aufgrund dieser Verbesserung zur besten Prognosegleichung geworden ist und die bisherige fitteste Gleichung abgelöst hat. Sie stellt nicht die Verbesserung der fittesten Gleichung zweier aufeinanderfolgenden Generationen dar, sondern kennzeichnet die prozentuale Verbesserung einer vormals schlechten Prognosegleichung zur Fittesten. So hat sich beispielsweise eine Prognosegleichung in Generation 13 aufgrund einer genetischen Operation um 46.6% verbessert und stellt infolgedessen die fitteste Prognosegleichung mit einem Out-of-Sample Vorhersagewert von 1.230 (Strategie A) dar.

Die Werte für die Building Block Analyse in Tabelle 4.28 (Strategie B) wurden dabei folgendermaßen berechnet: Zunächst wird für jeden Lauf die durchschnittlichen Fitneßänderungen aufgrund genetischer Operationen für jede Generation ausgerechnet. Anschließend wird für jeden Lauf diejenige Generation mit den dazugehörigen Schätzwerten ausgewählt, bei der die aggregierten Fitneßänderungen die 1% Marke unterschreiten. Die insgesamt 100 herausgefilterten Werte werden nach der Güte ihrer Richtungsprognosen aus der Trainingsphase geordnet. Die in Tabelle 4.28 dargestellten Ergebnisse geben die beste Prognosegleichung bzw. den Durchschnitt der besten 10 Prognosegleichungen wieder. Die so berechneten Ergebnisse schneiden zwar besser als ein Random Walk ab, aber die Prognosewerte weichen wiederum nicht signifikant von diesem ab. Betrachtet man die

Modell	Generation	In-Sample	Training Minimum	Prognose	Knoten	Lauf
beste Funktion ¹	13	0.777	0.136	1.031 (0.54)	42	24
besten 10 Funktionen ¹		0.821	0.140	1.155 (0.56)	38.2	1-100

Tabelle 4.28: Building Block; ¹ . Strategie B.

im Histogramm 4.11 abgebildeten 100 herausgefilterten Werte, so zeigt sich, daß der prozentuale Anteil schlechter Prognosewerte bei dieser Auswahlstrategie im Vergleich zu den anderen Strategien sehr hoch ausfällt.

Läßt man die bisherigen Untersuchungen Revue passieren, so hat sich gezeigt, daß die Prognosegüte innerhalb der ersten 200 Generationen besser gewesen sind als zwischen der 200. und 1000. Generation. Dies bedeutet, daß für die genetische Entwicklung von Prognosegleichungen die Suchintensivität nicht das ausschlaggebende Kriterium darstellt. Wenn demnach die Lösungskandidaten innerhalb der ersten 200 Generationen zu suchen sind, stellt sich die Frage, welche Generation innerhalb eines Laufs die vermeintlich besten Ergebnisse aufweist. Die Werte einer In-Sample MSE Schätzung werden stetig besser, so daß man keinen weiteren Anhaltspunkt für eine intelligente Auswahl bekommt. Die Punkt- als auch Richtungsprognosen der Trainingsphase liegen in der Regel nicht nur sehr dicht beieinander, sondern sie sind darüber hinaus sprunghaft in ihrer Entwicklung ohne erkennbares Muster. Diese Auswahlverfahren können die Frage, welche Generation man auswählen soll, nicht beantworten. Eine mögliche Lösung bietet die oben vorgestellte Kombination aus einer Building Block Analyse und der Richtungsprognose. Die Kombination bietet den Vorteil, daß die Wahl der Generation recht frühzeitig erfolgt, während die Richtungsprognose der Trainingsphase sicherstellt, daß von den anhand des ersten Kriteriums ausgewählten Prognosegleichungen erneut die besten selektiert werden. Allerdings zeigt das Histogramm, daß auch diese Methode kein zuverlässiges Auswahlverfahren darstellt.

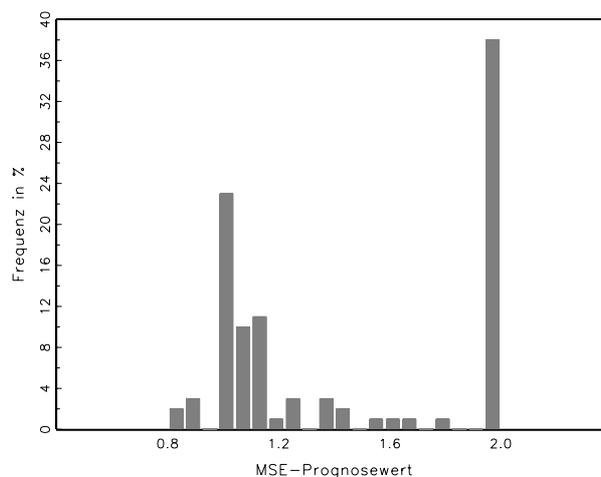


Abbildung 4.11: Building Block Analyse: Histogramm der Punktprognosen.

Modell	Generation	In-Sample	Training Minimum	Prognose	Knoten	Lauf
beste Funktion ¹	200	-0.318	0.134	1.024 (0.56)	94	88
beste Funktion ²	166	-0.218	0.133	1.023 (0.56)	92	88
besten 10 Funktionen ¹	200	-0.289	0.158	1.023 (0.56)	62.4	1-100
globales Minimum ³	34	-0.213	0.107	0.842 (0.64)	14	100
globale 10 Minima ³	1 bis 200	-0.255	0.112	0.850 (0.67)	29.4	1-100

Tabelle 4.29: Informationskriterium; ¹ aus In-Sample-Phase; ² aus Lauf 88: Funktion mit bestem Trainingwert; ³ aus Trainingsphase. modifizierte Diebold-Mariano Teststatistik für globales Minimum: -0.365.

Modell	Lauf	Generation	Trefferquote	Teststatistik	Signifikanz
beste Funktion ¹	88	200	50.704	0.006	0.94
beste Funktion ²	88	166	49.296	0.006	0.94
besten 10 Funktionen ¹	88	1 bis 200	49.296	0.006	0.94
globales Minimum ³	100	34	50.704	0.006	0.94
globale 10 Minima ³		1 bis 200	50.704	0.006	0.94

Tabelle 4.30: Informationskriterium; Richtungsprognose für Out-of-Sample Periode; Trefferquote in %. ¹ aus In-Sample-Phase; ² aus Lauf 88: Funktion mit bestem Trainingwert; ³ aus Trainingsphase.

4.4.5 In-Sample, Trainingsphase, Out-of-Sample und Informationskriterium

Eine andere methodische Herangehensweise, die bereits bei der Entwicklung der Prognosemodelle ansetzt, kann mit Hilfe eines Informationskriteriums erreicht werden. Das Kriterium lehnt sich an die Untersuchungen der Sensitivitätsanalyse in Kapitel 2 an und berechnet sich wie folgt:

$$AIC = \ln(MSE) + \frac{2 * m}{500} \quad (4.47)$$

mit m = Anzahl der Knoten. Die Gewichtung mit $\frac{1}{500}$ entspricht den Ergebnissen aus der Sensitivitätsanalyse und führte in der vorliegenden Simulation zu einer Reduzierung der durchschnittlichen Knotenanzahl auf 34.1 Knoten. Die Auswahl der Prognosegleichungen folgt den gleichen Strategien wie in den Abschnitten zuvor, wobei die Anzahl der Generationen erneut auf 200 beschränkt wurde. Die in Tabelle 4.29 aufgeführten Prognosen weisen zwar niedrigere Werte auf als die Vorhersagen eines Random Walks, aber sie weichen erneut nicht signifikant ab. Die Länge der Prognosegleichungen der besten Trainingsvorhersagen fallen mit 14 bzw. 29.4 Knoten relativ kurz aus, so daß leichte Parallelen zu den Untersuchungsergebnissen von Lütkepohl erkennbar sind, wonach sparsamer parametrisierte Regressionsgleichungen durch Verwenden entsprechender Informationskriterien tendenziell bessere Prognoseergebnisse erzielen. Unterstützt werden die Vermutungen durch die

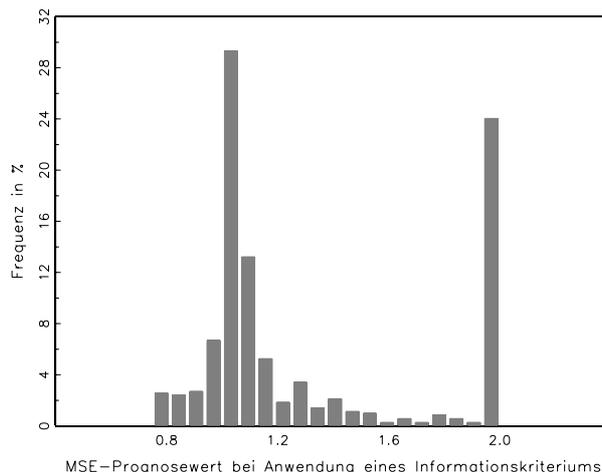


Abbildung 4.12: MSE Prognosewerte bei Verwendung eines Informationskriteriums.

in Abbildung 4.12 dargestellte Verteilung sämtlicher genetisch entwickelter Gleichungen zusammen mit ihren Out-of-Sample Vorhersagen. Im Vergleich zu den Untersuchungen, die auf einem MSE Fitneßkriterium basieren, generiert das Modell unter Verwendung eines Informationskriteriums eine wesentlich bessere Verteilung von Inflationsvorhersagen²⁷. Der Anteil schlechter Prognosegleichungen mit MSE-Werten > 2.0 beträgt nur ungefähr 25% im Gegensatz zu 35 – 45% der anderen Verteilungen und einem MSE als Fitneßkriterium. Darüber hinaus liegt der modale Wert bei ungefähr $MSE = 1.0$ und damit deutlich unterhalb der modalen Werte anderer Verteilungen sowie des Random Walks. Dieses ermutigende Ergebnis soll in einem neuen Simulationslauf überprüft werden, dessen Ergebnis im Anhang C.4 auf Seite 154 aufgeführt sind. Leider kann die zweite Simulation die Ergebnisse nicht bestätigen. Der Anteil schlechter Prognosegleichungen mit MSE-Werten > 2.0 beträgt beim MSE Kriterium nur ungefähr 21% im Gegensatz zu 27% bei Anwendung eines Informationskriteriums. Im Hinblick auf die Richtungsprognosen ergeben sich keine neuen Erkenntnisse.

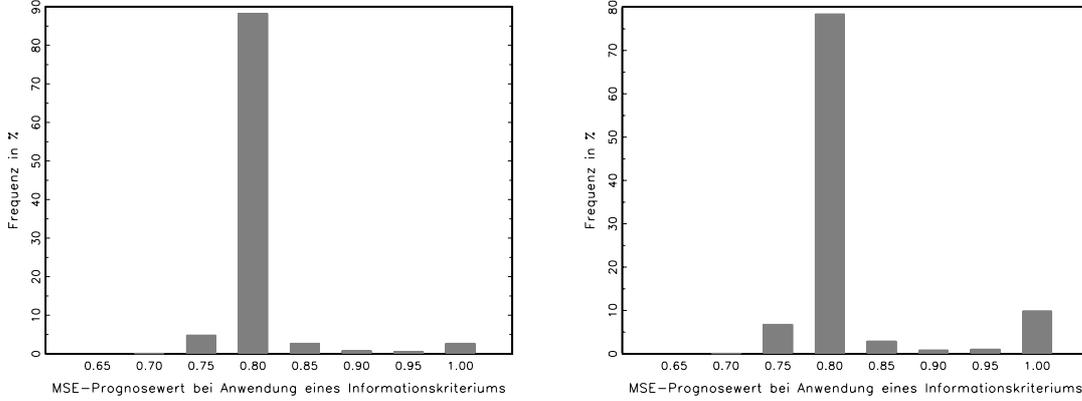
Vergleicht man die genetisch entwickelten Modelle mit den Ergebnissen des Vektorfehlerkorrekturmodells (VECM), so erzeugen die Genetischen Programme zwar bessere In-Sample Schätzungen, dafür schneiden die Richtungsprognosen des GP GAUSS schlechter ab. Während das VECM 54.93 % der Inflationsänderungen korrekt vorhersagt, liegt der Wert für Genetische Programme im Durchschnitt bei ungefähr 50 %.

4.4.6 Prognose von Inflationsänderungen

Zum Schluß dieses Kapitels sollen Inflationsänderungsraten anhand der Zinssätze i_t und ihren Änderungsraten Δi_t geschätzt werden. Die kurz- und langfristigen Zinssätze gehen mit bis zu 12 Zeitverzögerungen in die Terminalmenge von GP GAUSS ein. Dabei wird die Populationsgröße auf 100 Gleichungen und die Anzahl der Generationen auf 200 bei ins-

²⁷Vergleiche dazu die Histogramme unter Verwendung eines MSE Fitneßkriteriums (Anhang, S. 154), der Building Block Analyse 101 sowie der Backpropagation Analyse (Anhang, S. 154).

gesamt 100 Laufen festgelegt. Der Random Walk liefert als Benchmarkmodell einen MSE Prognosewert von 0.83 bei einer Trefferquote von 41.43% (Teststatistik = 1.036). Dabei werden die Ergebnisse erneut zwischen einem MSE- und einem Informationskriterium als Fitnekriterium miteinander verglichen. Dazu sollen als Erstes die beiden Histogramme



(a) Informationskriterium: Histogramm der Inflationsanderungsraten (b) MSE: Histogramm der Inflationsanderungsraten

Abbildung 4.13: Die Inflationsanderungsraten $\Delta\pi_t$ werden anhand der Zinssatze i_t und Δi_t geschatzt.

in Abbildung 4.13 betrachtet werden. Prognosewerte groer als $MSE = 1$ wurden auf den Wert eins gestutzt, um die Graphik anschaulicher zu gestalten. Es zeigt sich, da die modalen Werte unterhalb eines Random Walks liegen. Darber hinaus ist der Anteil schlechter Prognosen weitaus geringer als bei den geschatzten Inflationsraten der vorherigen Abschnitte.

Auch Tabelle 4.31 spiegelt diese Ergebnisse wieder. Alle Vorhersagewerte sind besser als ein Random Walk. Allerdings zeigt sich aufs Neue, da kein Vorhersagewert signifikant ist. Das Gleiche trifft fur die Richtungsprognosen zu. Die folgende genetisch entwickelte Gleichung gibt die ex-post beste Prognosegleichung wieder, die GP GAUSS auf Basis eines MSE Fitnekriteriums generiert hat:

$$\Delta\pi_t^{24} = (((\Delta i^2[t-3]) - (\cos((\Delta i^0[t-8]) + ((v) * (\sin(((60) - ((\Delta i^0[t-9]) - (((\Delta i^2[t-3]) - (\cos((\Delta i^0[t-8]) + ((v) * (\sin((i^2[t-10])) - (\text{abs}(17) * (\cos((\Delta i^0[t-8])))))))))) - (\cos((i^2[t-10])))) - ((27) + (i^2[t-9])))) - (\text{abs}(17) * (\cos((\Delta i^0[t-8])))))))))/11) \quad (4.48)$$

Abgesehen vom kurzfristigen Zinssatz in Niveaugroen i_t^0 finden sich alle Zeitreihen mit bis zu zehn Zeitverzogerungen in der besten ex-post Gleichung 4.48 wieder.

MSE Fitneßkriterium				
Modell	Out-of-Sample	Signifikanz	Trefferquote	Signifikanz
beste Funktion ¹	0.76	0.53	50.00	0.97
beste Funktion ²	0.76	0.53	50.00	0.97
besten 10 Funktionen ¹	0.76	0.53	50.00	0.97
globales Minimum ³	0.70	0.56	54.28	0.63
globale 10 Minima ³	0.68	0.58	51.42	0.87
AIC Fitneßkriterium				
Modell	Out-of-Sample	Signifikanz	Trefferquote	Signifikanz
beste Funktion ¹	0.76	0.53	50.00	0.97
beste Funktion ²	0.76	0.53	50.00	0.97
besten 10 Funktionen ¹	0.76	0.53	48.57	0.85
globales Minimum ³	0.76	0.53	48.57	0.85
globale 10 Minima ³	0.76	0.53	48.57	0.85
Random Walk	0.83		41.43	0.31

Tabelle 4.31: ¹ aus In-Sample-Phase; ² aus Lauf 88: Funktion mit bestem Trainingwert; ³ aus Trainingsphase.

4.5 Zusammenfassung

Das vorliegende Kapitel setzt sich mit der Prognosefähigkeit genetisch entwickelter Modelle auseinander und vergleicht diese mit populären, ökonomischen Modellen. Zwei Ziele standen dabei im Vordergrund der Untersuchungen: die Prognosegüte und die Optimierung der ex-anten Auswahl guter Prognosegleichungen, wobei der Schwerpunkt der Betrachtungen bei der Überprüfung verschiedener Auswahlstrategien lag. Als Vergleichsmodelle dienten neben dem naiven Modell eines Random Walks, die zwei weiteren univariaten ARIMA und ARFIMA Zeitreihenmodelle sowie das multivariate Vektorfehlerkorrekturmodell nach dem Johansen Verfahren. Die Aufgabe der Modelle war die Prognose deutscher Inflationsraten auf der Grundlage des Mishkin Ansatzes für den Zeitraum Januar 1997 bis Dezember 2002, wobei sowohl Punkt- als auch Richtungsprognosen berechnet wurden. Der Mishkin Ansatz postuliert, daß die Zinsdifferenz zweier Anleihen unterschiedlicher Laufzeit bei Annahme konstanter Realzinsen Informationen über die zukünftige Inflationsänderungsrate enthält. Die Untersuchungen umfaßten Null-Kupon-Anleihen mit Restlaufzeiten zwischen einem und fünf Jahren sowie einer zweijährigen durchschnittlichen Inflationszeitreihe, wobei ein dreimonatiger Geldmarktzinssatz zur Bildung der Zinsdifferenzen herangezogen wurde. Die VECM Modelle ergaben, daß eine Null-Kupon-Anleihe mit einer Restlaufzeit von zwei Jahren und durchschnittliche zweijährige Inflationsraten die mit Abstand besten Vorhersagen generierten. Aus diesem Grund und wegen der zeitaufwendigen GP GAUSS Simulationen beschränkten sich die anschließenden Untersuchungen genetischer Modelle auf diesen Datensatz. Bei den Parametereinstellungen des Genetischen Programms wurden die Ergebnisse der Sensitivitätsanalyse aus Kapitel 2 weitestgehend übernommen. Insgesamt zeigen die Ergebnisse der Vergleichsmodelle, daß der Random

Walk bei Punktprognosen und die ARIMA und ARFIMA Modelle bei den Richtungsprognosen die besten Werte erzielt haben.

Die durchgeführten Analysen genetischer Prognosemodelle basieren in den Abschnitten 4.4.2 und 4.4.1 auf einer Anzahl von jeweils 50 Millionen und in den Abschnitten 4.4.5, 4.4.4 und 4.4.3 auf 20 Millionen potentiellen Lösungsgleichungen pro Simulationslauf. Da nach jeder Generation nur die Ergebnisse der besten Gleichung gespeichert wurden, reduziert sich die Lösungsmenge auf eine Matrix mit 100.000 Prognosegleichungen, die für die Auswahlstrategien zur Verfügung standen. Auf dieser Grundlage lassen sich folgende Ergebnisse festhalten:

1. Bei den In-Sample Schätzungen generiert GP GAUSS deutlich bessere mittlere quadratische Fehlerabweichungen als die Referenzmodelle.
2. In einer ex-post Analyse wurde nachgewiesen, daß GP GAUSS bessere Vorhersagen als die Benchmarkmodelle erzielen kann. Allerdings sind die Abweichungen nicht signifikant.
3. Die Zweiteilung der Datenmenge in eine In-Sample und eine Out-of-Sample Periode und Selektion der Prognosegleichung anhand des kleinsten In-Sample MSE Wertes liefert zum Teil keine realistischen Werte.
4. Die Dreiteilung der Datenmenge in eine In-Sample-, Trainings- und Out-of-Sample Periode und Selektion der Prognosegleichung anhand des besten Prognosewertes aus der Trainingsphase aus dem Lauf mit dem minimalen In-Sample MSE Wert sowie die Selektion anhand der besten globalen Trainingsprognose erzeugen ebenfalls zum Teil sehr schlechte Inflationsvorhersagen.
5. Die Implementierung des Backpropagation Moduls konnte die In-Sample Schätzung nicht positiv beeinflussen bei gleichzeitiger Verschlechterung der Vorhersagegüte.
6. Bei der Building Block Analyse wurde für die Wahl der Prognosegleichung eine zwei-stufige Strategie angewendet: In einem ersten Schritt wurden für jeden Lauf und jede Generation die absoluten Fitneßänderungen sämtlicher Gleichungen aggregiert, um diejenige Prognosegleichung auszuwählen, bei der die Summe der Fitneßänderungen die 1%-Marke unterschreitet. Aus den 100 selektierten Gleichungen wurden schließlich diejenigen mit den besten Trainingsrichtungsprognosen verwendet. Auch diese kombinierte Auswahlstrategie führte zu keiner Verbesserung der Ergebnisse.
7. Die auf den ersten Blick besten Ergebnisse erzielten Genetische Programme, die ein Informationskriterium als Fitneßkriterium verwendet haben. Der Vergleich verschiedener Histogramme zeigte zunächst, daß das Verwenden eines Informationskriteriums die Prognosequalität genetisch entwickelter Gleichungen sichtbar verbessern kann. Die Ergebnisse konnten bei einem wiederholten Simulationslauf aber nicht bestätigt werden.
8. Die Schätzung von Inflationsänderungsraten hat gezeigt, daß GP GAUSS zwar bessere Punkt- und Richtungsprognosen erzeugt als ein Random Walk, aber die Vorhersagen sind nicht signifikant.

Bei den Untersuchungen konnte insgesamt festgestellt werden, daß weder Punktprognosen noch Richtungsprognosen, die bessere Werte als ein Random Walk aufgewiesen haben, signifikant sind. In Bezug auf die Richtungsprognosen konnten die genetischen Modelle nur selten die Werte der ARIMA und ARFIMA Modelle annähernd erreichen. Wenngleich die Ergebnisse beim Vergleich zwischen einem MSE und Informationskriterium durchwachsen sind, sollen sie im nächsten Kapitel erneut miteinander verglichen werden. Das folgende Kapitel setzt sich mit Wechselkursmodellen auseinander, deren Prognosemodelle im Unterschied zu den hier erstellten Modellen rekursiv erfolgt.

Kapitel 5

Wechselkursprognosemodelle

Nachdem im vorherigen Kapitel die Untersuchung verschiedener Auswahlstrategien den Schwerpunkt der Analyse bildete, werden im vorliegenden Kapitel Wechselkursprognosen auf der Grundlage geldpolitischer Ansätze erstellt, die in den 1970er und 1980er Jahren nach dem Zusammenbruch des Bretton-Woods Systems konzipiert worden sind. Im Einzelnen handelt es sich um die beiden noch heute populären monetären Ansätze des flexiblen Preismodells von Frenkel [45] (1976) und Bilson [8] (1979), dem monetären Modell mit trägen Preisen von Dornbusch und Frankel [30] (1976) sowie um einen portfoliotheoretischen Ansatz. Dabei lehnen sich die Untersuchungen an die Artikel von Plasmans et al. [104] und Verkooijen et al. [127] an, die mit Hilfe künstlicher neuronaler Netze mögliche nichtlineare Elemente in den Finanzmarktzeitreihen für die Prognose auszunutzen versuchten. Gute Wechselkursprognosen sind sowohl auf Zentralbank- als auch auf Unternehmensebene wünschenswert, da Wechselkurschwankungen zum einen über ausländische Güter im Inland einen Teil der inländischen Inflationsrate und des Realeinkommens mitbestimmen können und zum anderen Einfluß auf die internationale Wettbewerbsfähigkeit von Unternehmen ausüben. Die Effekte der wirtschaftspolitischen Zielgrößen steigen mit dem Offenheitsgrad einer Volkswirtschaft, d.h. mit der Abnahme nationaler Handelsbarrieren gegenüber dem Ausland¹. Die Konzeption der Prognosemodelle ist in diesem Kapitel gegenüber den Inflationsprognosemodellen etwas anders aufgebaut. Bisher erfolgten die Schätzungen und Prognosen der genetisch entwickelten wie auch der Vergleichsmodelle anhand einer fixen In- und Out-of-Sample Datenmenge. Dies ändert sich im Folgenden durch das Anwenden einer rekursiven Regressionsmethode. Jeder Beobachtungswert der Out-of-Sample Menge, der sich im Zeitablauf von einem unbekanntem zu einem realisierten Wert wandelt, wird in der Folgeperiode der In-Sample Menge hinzugefügt und die Modelle werden neu geschätzt. Dadurch vergrößert sich die In-Sample Menge in jeder Periode um einen Wert, während gleichzeitig die Anzahl der zu prognostizierenden Out-of-Sample Werte sich entsprechend verkleinert. Um diesen Aspekt im Genetischen Programm zu berücksichtigen, erhöht sich mit jedem neuen Lauf die In-Sample Menge um jeweils einen Wert bei gleichzeitiger Verringerung der Out-of-Sample Menge. Folglich stehen für die Auswahl einer Prognosegleichung nur die in *einem* Lauf generierten Lösungskandidaten

¹Vgl. auch Willms [130] (1992).

und nicht mehr sämtliche produzierten Gleichungen zur Verfügung, wobei die Auswahl anhand der besten In-Sample Schätzung nach 200 Generationen vorgenommen wird.

Mit der Untersuchung von Wechselkursprognosen werden mehrere Ziele verfolgt: erstens soll mit dem Anwenden des Genetischen Programmierens überprüft werden, ob die Anwendung unterschiedlicher Suchalgorithmen auf die gleiche Problemstellung qualitativ unterschiedliche Resultate hervorruft. Des Weiteren wird erneut das MSE und das Informationskriterium miteinander verglichen. Drittens sind die Untersuchungen wesentlich umfangreicher als die von Plasmans et al. und Verkooijens et al., da neben den beiden verschiedenen Fitneßkriterien unterschiedliche Datenzeiträume betrachtet und neben Punktprognosen erneut Richtungsvorhersagen berechnet werden.

Das Kapitel beginnt im Abschnitt 5.1 mit einer kurzen Erläuterung der verschiedenen geldpolitischen Ansätze und der Herleitung der zugrundeliegenden Wechselkursbestimmungsgleichungen².

5.1 Theoretische Ansätze zur Bestimmung des Wechselkursverhaltens

Da ein wesentlicher Bestandteil der monetären Ansätze die Kaufkraft- und die Zinsparitätentheorie bilden, soll eine kurze Erläuterung dieser Theorien den anderen Modellen vorangestellt werden. Der Hauptgedanke der Kaufkraftparität ist die proportionale Anpassung zwischen den Wechselkursen und den Konsumentenpreisen, um die Kaufkraft der verschiedenen Währungen aufrecht zu erhalten. Die strikte Version postuliert, daß der reale Wechselkurs zu jedem Zeitpunkt in allen Staaten der gleiche ist. Dagegen fordert die Zinsparitätentheorie, daß die Zinsdifferenz zweier Staaten der erwarteten Änderungsrate des Wechselkurses entsprechen muß.

5.1.1 Kaufkraftparitätentheorie

Der älteste Ansatz in der monetären Außenwirtschaftstheorie zur Erklärung von Wechselkursschwankungen stellt die Kaufkraftparität dar. Dogmengeschichtlich wird der Ansatz dem Ökonomen Cassel (1918) zugeschrieben³, wengleich erste Ansätze bis ins 16. Jahrhundert zurückreichen. Die Kaufkraftparitätentheorie besagt in ihrer ursprünglichen Form, daß der Wechselkurs als Regulator zwischen in- und ausländischen Gütermarktpreisen fungiert und für einen Ausgleich der Kaufkraft zweier Währungen sorgt. Demnach bewirken der internationale Handel und Wettbewerb, daß ein bestimmtes Güterbündel nach Umrechnung mit dem Wechselkurs in allen Ländern einen einheitlichen Preis erzielt. Die in der Realität zu beobachtenden Abweichungen der Wechselkurse von der Kaufkraftparitätentheorie und damit Abweichungen vom 'Gesetz des einheitlichen Preises' werden darauf zurückgeführt, daß nur ein Teil der in der Volkswirtschaft hergestellten und konsumierten Güter international handelbar ist, und daß institutionelle Restriktionen wie Transportkosten, Importkontingentierungen, Zölle und Steuern Preisunterschiede verursachen.

²Siehe auch Gärtner [46], Mark [86] oder Pilbeam [103].

³Siehe auch Bofinger [11].

Für die formale Herleitung wird eine in- und eine ausländische Volkswirtschaft betrachtet⁴, deren Preise P und P^* sich jeweils anhand eines Güterbündels bestehend aus n handelbaren Gütern berechnen: $P := \sum_{k=1}^n \alpha_k p_k$ und $P^* := \sum_{k=1}^n \alpha_k p_k^*$ mit p_k als Preis des in- und ausländischen Gutes k und α_k als Gewichtungssparameter des Gutes k innerhalb des Preisindexes. Unter den Annahmen homogener Güter und der Abwesenheit institutioneller Restriktionen muß aufgrund von Arbitragefreiheit das „Gesetz des einheitlichen Preises“ gelten, so daß die Kaufkraft einer Währung - gemessen durch die Menge eines Gutes, die man mit einer Geldeinheit erwerben kann - im Vergleich zu einer anderen Währung über den Wechselkurs ausgeglichen wird. Somit bestimmt sich der logarithmierte Wechselkurs über die Differenz der logarithmierten in- und ausländischen Preise und stellt in dieser Form die absolute Kaufkraftparität dar:

$$S_t = \frac{P_t}{P_t^*} \Rightarrow \log S_t = \log P_t - \log P_t^* \Rightarrow s_t = p_t - p_t^* \quad (5.1)$$

mit S_t als nominalem Kassawechselkurs. Eine Transformation dieser Gleichung in Änderungsraten führt zur relativen Kaufkraftparität:

$$s_t - s_{t-1} = (p_t - p_{t-1}) - (p_t^* - p_{t-1}^*) \Rightarrow \Delta s_t = \pi_t - \pi_t^* \quad (5.2)$$

In ihrer modifizierten Version besagt die Kaufkraftparitätentheorie, daß die Änderungsrate des Wechselkurses der Differenz der Inflationsraten zwischen In- und Ausland entspricht. Eine höhere (niedrigere) inländische Inflationsrate π , die die Wettbewerbsfähigkeit ausländischer Güter verbessert (verschlechtert), muß über eine sofortige Abwertung (Aufwertung) des Wechselkurses $s_t \uparrow$ ($s_t \downarrow$) wieder ausgeglichen werden. Bei Gültigkeit der Kaufkraftparität muß folglich auch der reale Wechselkurs konstant bleiben.

5.1.2 Zinsparitätentheorie

Während sich der Wechselkurs bei der Kaufkraftparitätentheorie durch unterschiedliche Preisniveaus im In- und Ausland bestimmt, postuliert die Zinsparitätentheorie, daß sich die Differenz zwischen Devisenkassa- und -terminkurs durch die Zinsdifferenz homogener in- und ausländischer Wertpapiere erklären läßt. Ihre Herleitung ist simple und geht laut Baille & McMahon [4] auf Walther Lotz (1884) zurück. Durch Anlage eines Betrages A in ein in- und ausländisches Wertpapier und gleichzeitiger Absicherung der eingegangenen Fremdwährungsposition $A \frac{1}{s_t}$ durch ein Termingeschäft s_t^T in gleicher Höhe kann die Zinsparitätentheorie folgendermaßen formuliert werden:

$$A(1+i) = A \frac{s_t^T}{s_t} (1+i_t^*) \Rightarrow \frac{s_t^T}{s_t} = \frac{1+i_t}{1+i_t^*} \Rightarrow \frac{s_t^T}{s_t} - 1 = \frac{1+i_t}{1+i_t^*} - 1 \Rightarrow \frac{s_t^T - s_t}{s_t} \approx i_t - i_t^* \quad (5.3)$$

Bei Verletzung dieser Gleichung könnte ein Investor die Differenz durch gewinnbringende Arbitragegeschäfte solange ausnutzen, bis die Differenz aufgrund der Kapitalbewegungen auf den Finanzmärkten verschwunden ist. Die aufgrund des Termingeschäfts als gedeckte oder gesicherte bezeichnete Zinsparität muß zu jedem Zeitpunkt erfüllt sein. Erwartet der

⁴Die ausländischen Variablen sind mit einem * gekennzeichnet.

Händler in Zukunft einen günstigeren Wechselkurs als das garantierte Umtauschverhältnis des Termingeschäfts, könnte es für den Händler attraktiver sein abzuwarten und das Wechselkursrisiko in Kauf zu nehmen. In diesem Fall müssen die Akteure auf dem Terminmarkt eine Risikoprämie bezahlen, die der Differenz zwischen dem Termingeschäft und dem zukünftigen Kassakurs entspricht. Existiert am Devisenmarkt keine Risikoprämie, stimmt also der erwartete zukünftige Wechselkurs mit dem Termingeschäft überein, impliziert die gedeckte Zinsparität das Vorliegen der ungedeckten Zinsparität (UIP). Formal betrachtet, wird bei der UIP die abgesicherte Devisenposition s_t^T durch die subjektiven Erwartungen der Anleger $E(s_{t+k})$ ersetzt:

$$E(s_{t+k}) - s_t = i_t - i_t^* \quad (5.4)$$

mit $E(s_{t+k})$ als erwartetem Wechselkurs zum Zeitpunkt $t + k$. Die UIP fußt auf den Annahmen perfekter Substituierbarkeit zwischen in- und ausländischen Wertpapieren und auf uneingeschränkter Kapitalmobilität. Darüber hinaus werden vollkommen effiziente Finanzmärkte angenommen, so daß beispielsweise Transaktionskosten, unterschiedliche Steuersysteme oder Risikoprämien bei Termingeschäften unberücksichtigt bleiben.

5.1.3 Monetäre Modelle der Wechselkursentwicklung

Die beiden monetären Ansätze des flexiblen und rigiden Preismodells und der Portfolioansatz erklären die Wechselkursbewegungen über Bestandsanpassungsprozesse auf den Vermögenmärkten. Demnach ist nicht mehr der relative Güterpreis sondern vorwiegend der relative Preis zwischen in- und ausländischen Finanzaktiva, bewertet in der jeweiligen Währung, entscheidend. Folglich stellt sich ein Gleichgewichtswechselkurs ein, sobald die in- und ausländischen Investoren keine Vermögensumschichtungen mehr vornehmen und ein bestimmter Bestand an Finanzaktiva existiert. Allen Finanzmarktmodellen ist die Annahme perfekter Kapitalmobilität gemein, so daß aufgrund von Arbitrageüberlegungen die gedeckte Zinsparität gelten muß. Der zentrale Unterschied besteht bei der Annahme der Substituierbarkeit in- und ausländischer Wertpapiere, die die Dichotomisierung der Wechselkursmodelle in monetäre und portfoliotheoretische Ansätze begründet. Demzufolge treffen das flexible und rigide Preismodell die Annahme homogener in- und ausländischer Wertpapiere mit der Konsequenz perfekter Substituierbarkeit und der Gültigkeit der ungedeckten Zinsparität. Da die risikoneutralen Investoren die in- und ausländischen Wertpapiere in ihrem Risiko nicht unterscheiden, ist der Terminkurs ein unverzerrter Schätzer für den zukünftigen Devisenkurs. Entspräche die Zinsdifferenz zwischen dem In- und Ausland nicht den erwarteten Änderungen auf dem Kassamarkt, könnten die Anleger Arbitragegewinne erzielen. Die Annahme eines homogenen in- und ausländischen Wertpapiermarktes führt zu einer weiteren entscheidenden Modellimplikation gegenüber dem Portfolioansatz: die Wechselkursentwicklung wird bei den monetären Ansätzen allein auf den Geldmärkten bestimmt. Dagegen setzt der Portfolioansatz eine imperfekte Substituierbarkeit der Wertpapiere voraus, so daß ein Anleger ausländische Finanzaktiva kauft, wenn der erwartete Ertrag die Verzinsung aus einer inländischen Anlage inklusive einer Risikoprämie übersteigt, die er für das Wechselkursrisiko verlangt. Folglich sind die Bedingungen für die Gültigkeit einer ungedeckten Zinsparität nicht erfüllt und die Wechselkursbewegungen

werden neben dem Geldmarkt auch durch Bestandsänderungen auf dem Kapitalmarkt beeinflußt.

5.1.4 Monetäre Theorie bei flexiblen Preisen

Die Bestimmungsgleichung des Wechselkurses leitet sich bei den monetären Ansätzen direkt aus der absoluten Kaufkraftparität her (Gleichung 5.1), die aufgrund der Annahme vollkommen flexibler Gütermarktpreise p eine permante Gültigkeit besitzt⁵:

$$s_t = p_t - p_t^* \quad (5.5)$$

Wegen der bereits oben angeführten Annahmen vollkommener Kapitalmobilität und perfekter Substituierbarkeit von Finanzaktiva, die einen einheitlichen, stets im Gleichgewicht befindlichen, internationalen Kapitalmarkt bewirken, wird die Geldnachfragefunktion zur entscheidenden Bestimmungsgleichung des Modells:

$$m_t^d = p_t + \eta y_t - \lambda i_t \quad (5.6)$$

mit m_t^d als Geldnachfrage, y_t als Realeinkommen und i_t als Nominalzinssatz. Da alle Variablen bis auf den Zinssatz in Logarithmen ausgedrückt werden, handelt es sich bei $\eta = (\Delta M_t / M_t) / (\Delta Y_t / Y_t)$ um die Einkommenselastizität und bei $\lambda = (\Delta M_t / M_t) / \Delta i_t$ um die Semi-Zinselastizität der Geldnachfrage. Die Preisniveaus im In- und Ausland sind durch ein Gleichgewicht auf den nationalen Geldmärkten bestimmt, so daß mit einem exogen festgelegten Geldangebot m_t die nachfolgende Gleichung die Bedingung für ein Geldmarktgleichgewicht widerspiegelt:

$$m_t = p_t + \eta y_t - \lambda i_t \quad (5.7)$$

Analog dazu verhält sich das Gleichgewicht auf dem ausländischen Geldmarkt:

$$m_t^* = p_t^* + \eta^* y_t^* - \lambda^* i_t^* \quad (5.8)$$

Werden zusätzlich identische in- und ausländische Elastizitäten angenommen, erhält man durch Subtrahieren der beiden Gleichungen 5.7 und 5.8

$$m_t - m_t^* = p_t - p_t^* + \eta(y_t - y_t^*) - \sigma(i_t - i_t^*) \quad (5.9)$$

sowie durch Auflösen der Gleichung nach $p_t - p_t^*$ und anschließendem Einsetzen in die KKP-Hypothese 5.5 die monetäre Bestimmungsgleichung des Wechselkurses mit flexiblen Gütermarktpreisen:

$$s_t = (m_t - m_t^*) - \eta(y_t - y_t^*) + \sigma(i_t - i_t^*) \quad (5.10)$$

Demnach bestimmt sich der Wechselkurs aus der Differenz der Geldbestände und der Einkommensniveaus zwischen dem In- und Ausland sowie der Zinsdifferenz.

⁵Eine Erhöhung des inländischen Geldangebots verursacht eine simultane Erhöhung der inländischen Gütermarktpreise und Abwertung der inländischen Währung um den gleichen Prozentsatz, so daß neben der permanenten Gültigkeit der KKP eine konstante reale Wechselkursrate vorliegt.

Führt die inländische Zentralbank beispielsweise eine expansive Geldpolitik mit einem Angebotsüberschuß auf dem Geldmarkt durch, dann versuchen die Wirtschaftssubjekte die vorliegende überschüssige Realkasse durch Güterkäufe wieder abzubauen. Aufgrund eines gegebenen Vollbeschäftigungsoutput auf dem Gütermarkt und vollkommen flexibler Preise erfolgt eine sofortige Erhöhung des Preisniveaus mit einem neuen simultanen Gleichgewicht auf dem Geld- und Gütermarkt. Somit werden die Aussagen der Quantitätstheorie bestätigt, nach der eine monetäre Expansion keine realwirtschaftlichen Effekte auslösen kann (Neutralität des Geldes). Der Anstieg des inländischen Gütermarktpreises führt über die Kaufkraftparität zu einer Abwertung der inländischen Währung, so daß der reale Wechselkurs aufrecht erhalten bleibt.

Die Hauptkritikpunkte des flexiblen Preismodells bilden die beiden zentralen Hypothesen der ungedeckten Zinsparität und der Annahme einer kontinuierlichen Gültigkeit der Kaufkraftparität sowohl im kurz- als auch langfristigen Zeitraum.

5.1.5 Monetäre Theorie bei kurzfristig trägen Preisen

Die letztgenannte Kritik wird im monetären Modell mit trägen Preisen, das auf einen Beitrag von Dornbusch [30] zurückgeht, aufgegriffen, in dem unterschiedliche Anpassungsgeschwindigkeiten auf den Finanz- und Gütermärkten zugelassen werden. Während auf den homogenen Finanzmärkten exogene Schocks aufgrund unbeschränkter Kapitalmobilität nach wie vor ohne Zeitverzögerung ausgeglichen werden, ist die vollständige Preisflexibilität auf den Gütermärkten aufgehoben. Kurzfristige Preisstarrheiten können zu temporären Ungleichgewichten führen, die sich langfristig wieder auflösen und in der Gültigkeit der Kaufkraftparität enden. Die in der langen Frist erhaltenen Preise stellen die fundamentalen Gleichgewichte auf den Gütermärkten und auf dem Devisenmarkt dar. Kurzfristig weicht der Wechselkurs vom fundamentalen Wechselkurs der Kaufkraftparität \bar{s}_t ab und wird formal durch den Term $\theta(\bar{s}_t - s_t)$ in die Gleichung der relativen Kaufkraftparität 5.2 integriert:

$$E(\Delta s_t) = \theta(\bar{s}_t - s_t) + \pi_t - \pi_t^* \quad (5.11)$$

mit θ als Preisanpassungsgeschwindigkeit. Langfristig entspricht der aktuelle Wechselkurs s_t dem fundamentalen Wechselkurs \bar{s} und es stellt sich die relative Kaufkraftparität 5.2 wieder ein. Durch Gleichsetzen von Gleichung 5.11 und der ungedeckten Zinsparität $E(\Delta s_t) = i_t - i_t^*$ resultiert die Gleichung

$$s_t - \bar{s} = -\frac{1}{\theta}(i_t - i_t^* - (\pi_t - \pi_t^*)) \quad (5.12)$$

Da langfristig das rigide und flexible Preismodell übereinstimmen, drückt die Wechselkursbestimmungsgleichung des flexiblen Preismodells den fundamentalen Wechselkurs aus:

$$\bar{s} = (m_t - m_t^*) - \eta(y_t - y_t^*) + \sigma(\pi_t - \pi_t^*) \quad (5.13)$$

Durch Einsetzen der Gleichung 5.13 in 5.12 resultiert schließlich die Wechselkursgleichung des monetären Modells mit rigiden Preisen:

$$s_t = m_t - m_t^* - \eta(y_t - y_t^*) - \frac{1}{\theta}(i_t - i_t^*) + \left(\sigma + \frac{1}{\theta}\right)(\pi_t - \pi_t^*) \quad (5.14)$$

Da die Gütermarktpreise langsamer zum langfristigen Gleichgewicht konvergieren als die Wertpapierpreise, weisen letztere eine stärkere Preisdynamik auf als beim flexiblen Preismodell, damit die Märkte sich vorübergehend in einem Gleichgewicht befinden. Der reale Wechselkurs weicht in der kurzen Frist im Gegensatz zum flexiblen Preismodell von seinem langfristigen Gleichgewicht ab und überschießt wie auch der nominale Wechselkurs in der kurzen und mittleren Frist die langfristig zu erwartenden Änderungen.

5.1.6 Portfoliotheorie

Sowohl das flexible als auch das rigide Preismodell enthalten zwei wesentliche Schwachstellen, die der Portfolioansatz durch die Hypothese einer unvollkommenen Substituierbarkeit in- und ausländischer Wertpapiere beseitigt. Die beiden Kritikpunkte umfassen die Vernachlässigung möglicher Handelsbilanzungleichgewichte zwischen zwei Ländern sowie die Annahme einheitlicher internationaler Kapitalmärkte. Liegt beispielsweise ein Handelsüberschuß vor, dann impliziert die aktuelle Handelsbilanz eine Akkumulation ausländischer Finanzaktiva durch inländische Wirtschaftssubjekte und somit eine Zunahme des inländischen Vermögens. Die Zunahme des Vermögens steigert wiederum die Konsumausgaben und damit verbunden die Geldnachfrage, die ihrerseits Einfluß auf den Wechselkurs ausübt. Zudem entfällt durch die Annahme einer unvollkommenen Substituierbarkeit in- und ausländischer Wertpapiere die Gültigkeit der ungedeckten Zinsparität. Dies hat zur Folge, daß die Wirtschaftssubjekte Risikoprämien für das Wechselkursrisiko verlangen und ihr Vermögen aufgrund eines heterogenen Kapitalmarktes entsprechend portfoliotheoretischer Überlegungen diversifizieren müssen. Das Netto-Finanzvermögen des privaten Sektors kann in drei Komponenten aufgeteilt werden: die nominale inländische Geldmenge, die inländischen Staatsanleihen und die durch inländische Investoren gehaltenen und in fremder Währung notierten ausländischen Staatsanleihen. Bei zugrundegelegten flexiblen Wechselkursen entspricht der Betrag eines Handelsbilanzüberschusses dem Kapitaldefizit. Die dadurch notwendige Risikoprämie wird bei der ungedeckten Zinsparität eingebaut:

$$i_t - i_t^* = E(\Delta s_t) + RP(b - f - s)_t \quad (5.15)$$

Durch eine linearisierte Version

$$\alpha + \beta(i_t - i_t^*) - E(\Delta s_t) = b - f - s \quad (5.16)$$

und Einsetzen in Gleichung 5.2 führt zu

$$s_t - \bar{s} = \frac{b - s - f - \alpha}{\beta\theta} - \frac{1}{\theta}((i_t - \pi_t^e) - (i_t^* - \pi_t^{e*})) \quad (5.17)$$

Zusammen mit Gleichung 5.13 stellt sich für den langfristigen Wert \bar{s} die Wechselkursgleichung der Portfoliotheorie ein:

$$s_t = \frac{-\alpha}{\theta\beta + 1} + \frac{1}{1 + 1/\theta\beta}(m_t - m_t^*) - \frac{\eta}{1 + 1/\theta\beta}(y_t - y_t^*) - \frac{1}{\theta + 1/\beta}(i_t - i_t^*) + \frac{\sigma + 1/\theta}{1 + 1/\theta\beta}(\pi_t^e - \pi_t^{e*}) + \frac{1}{\theta\beta + 1}(b - f) \quad (5.18)$$

	Japan	Deutschland	UK	Niederlande
Wechselkurs	I(1)	I(1)	I(1)	I(1)
Nominalzins	I(0)	I(1)	I(0)	I(0)
Geldmenge	I(1)	I(1)	I(1)	I(1)
Industrie Prod.	I(1)	I(1)	I(1)	I(1)
Inflation	I(0)	I(0)	I(0)	I(0)
Handelsbilanz	I(1)	I(1)	I(1)	I(1)

Tabelle 5.1: Test auf Stationarität anhand eines Unit Root Tests

5.2 Daten

Die oben hergeleiteten Wechselkursbestimmungsgleichungen des flexiblen und rigiden Preismodells sowie des Portfolioansatzes werden im Folgenden auf die Zeitreihen für Japan, Großbritannien, Niederlande und Deutschland angewendet mit den USA als Bezugsgröße⁶. Bei den monatlichen Zeitreihen⁷ handelt es sich um bilaterale Wechselkursraten S_t , Industrieproduktionsindizes IP_t , Konsumentenpreisindizes CP_t , Nominalzinssätze⁸ i_t , Geldmengenaggregate $M1$ oder $M0$ sowie um Handelsbilanzwerte TB_t und wurden den Datenbanken *Datastream* und *Main Economic Indicators* der OECD entnommen. Die Daten erstrecken sich über den Zeitraum Januar 1974 bis Oktober 2003 für Japan und Großbritannien bzw. bis Dezember 1998 für die beiden EU-Staaten Niederlande und Deutschland, wobei neben der Berücksichtigung des gesamten Zeitraumes ein zweiter Datensatz beginnend ab dem Louvre Abkommen untersucht wird. Die Inflationszeitreihen berechnen sich aus der Differenz der logarithmierten Konsumentenpreisindizes:

$$\pi_t = \ln(CP_t) - \ln(CP_{t-1}) \quad (5.19)$$

5.3 Zeitreiheneigenschaften

Wie bereits beim Mishkin Ansatz werden als Erstes die Zeitreihen mit Hilfe des ADF Tests auf Stationarität überprüft. Genau genommen, werden die Differenzen der logarithmierten Zeitreihenwerte zwischen den 'inländischen' Werten Japan, Deutschland, Großbritannien und Niederlande und den 'ausländischen' amerikanischen Daten berechnet mit Ausnahme der Wechselkurse und der Handelsbilanzwerte. Die Ergebnisse der Tabelle 5.1 zeigen, daß alle Zeitreihen mit Ausnahme der Inflationszeitreihen und der meisten Nominalzinsen (außer Deutschland) nichtstationär sind. Damit weichen die von Plasmans et

⁶Da bereits zu Beginn dieser Arbeit die empirischen Ergebnisse verschiedener Studien vorgestellt wurden, wird in diesem Zusammenhang auf Abschnitt 1.5 verwiesen.

⁷Siehe auch Anhang D.1. Die Zeitreihen orientieren sich an den beiden Artikeln von Plasmans et al. sowie Verkooijen et al., wobei für die Niederlande der deutsche Interbankenzinssatz FIBOR verwendet wird. Der niederländische Interbankenzinssatz AIBOR und der europäische EURIBOR existieren nicht für den betrachteten Zeitraum.

⁸kurzfristige Zinssätze: Interbankenzinssätze mit einer Restlaufzeit von 3 Monaten, kurzfristige festverzinsliche Wertpapiere. Für den gemeinsam vorhandenen Zeitraum beträgt der Korrelationskoeffizient zwischen FIBOR und AIBOR über 96%.

al. bzw. Verkooijen et al. kalkulierten Testergebnisse vor allem bei den Industrieproduktionsindices, den Inflations- und Zinszeitreihen ab, denen Beobachtungswerte bis Juli 1994 zur Verfügung standen. Die im Anhang D.1 aufgeführten Graphiken zeigen die teilweise gravierenden Änderungen im Zeitablauf nach Juli 1994, die die Abweichungen der Ergebnisse erklären können. Demnach kann für die Inflationsraten wie auch den Zinssätzen ein deutlicher Rückgang verzeichnet werden. Die Industrieproduktion spiegelt hingegen die konjunkturelle Aufschwungphase wider.

5.4 Lang- und kurzfristige rekursiv geschätzte Modelle

In den kommenden Abschnitten soll die Vorhersagegüte der verschiedenen Wechselkursmodelle für langfristige als auch kurzfristige Prognosehorizonte überprüft werden, deren ökonomischen Fundamentalvariablen in der Gleichung 5.20 nocheinmal zusammenfassend dargestellt sind:

$$s_t = f(r_t - r_t^*, m_t - m_t^*, ip_t - ip_t^*, \pi_t - \pi_t^*, TB_t, TB_t^*) + \epsilon \quad (5.20)$$

mit s_t als logarithmierten Wechselkurs, $r_t - r_t^*$ als Differenz des in- und ausländischen kurzfristigen Nominalzinssatzes, $m_t - m_t^*$ als Differenz der logarithmierten in- und ausländischen Geldmengen, $ip_t - ip_t^*$ als Differenz der logarithmierten Industrieproduktion, $\pi_t - \pi_t^*$ als Differenz der Inflationsrate und TB, TB^* als kumulierte Handelsbilanz und ϵ schließlich als Störterm. Die mit einem '*' gekennzeichneten Variablen indizieren die amerikanischen Zeitreihen. Während das flexible Preismodell die ersten drei Terme $r_t - r_t^*, m_t - m_t^*$ und $ip_t - ip_t^*$ berücksichtigt, schließt das rigide Preismodell die Differenz der Inflationsraten $\pi_t - \pi_t^*$ mit ein, wohingegen der Portfolioansatz sich auf $m_t - m_t^*$ und TB, TB^* beschränkt. Um mögliche langfristige Beziehungen zwischen den Wechselkurszeitreihen und den hier verwendeten fundamentalen Finanzmarktzeitreihen offenzulegen, werden die Modelle anhand dieser Niveauvariablen mittels einer rekursiven Regressionsmethode geschätzt und k-step-ahead Prognosen für die Monate $k = 1, 6, 12, 18$ und 24 berechnet. Die kurzfristigen Modelle benutzen hingegen die einmal differenzierten Niveauvariablen und schätzen mit Hilfe dieser Änderungsraten eine Ein-Monats-Prognose der Wechselkursänderungsrate. Als Gütemaß der einzelnen Prognosewerte wird der Root Mean Squared Error (RMSE) verwendet in Anlehnung an die Artikel von Meese & Rogoff sowie Plasmans et al.:

$$RMSE(k) = \sqrt{\sum_{t=n_0}^{N-k} (\hat{s}_{t+k} - s_{t+k})^2 / (N - k)} \quad (5.21)$$

mit k als Prognosehorizont, \hat{s} als geschätztem und s als realisiertem Wechselkurs. Die kontinuierliche Vergrößerung der In-Sample Menge um einen Wert und der anschließenden Schätzung sowie die Prognosen für die sich stetig verringernde Out-of-Sample Menge wiederholt sich solange, bis die In-Sample Menge der gesamten Menge abzüglich des Prognosehorizonts entspricht. Ein wichtiger Punkt bei der Erstellung der Prognosen bleibt anzumerken. Betrachtet man die hergeleiteten Schätzgleichungen genau, so fällt auf, daß die erklärenden Variablen gleichzeitig mit der vorherzusagenden Wechselkursvariablen vorliegen müssen. Dies bedeutet, daß im Grunde genommen auch die erklärenden - zu dem

Zeitpunkt nicht bekannten - Variablen prognostiziert werden müßten. Meese & Rogoff aber auch Plasmans et al. und Verkooijen et al. lösen diese Problematik, in dem sie die realisierten Werte einsetzen. Folglich stellen die Prognosen keine 'wahren' Prognosen dar. GP GAUSS schließt sich dieser Vorgehensweise an.

Die Untersuchungen von Plasmans et al. haben mehrere Ergebnisse offengelegt. Die Schätzung langfristiger Modelle auf Basis künstlicher neuronaler Netze konnten keine bessere Prognosegüte im Vergleich zu einem Random Walk erzielen, so daß keine Unterstützung für langfristige Beziehungen zwischen Wechselkursraten und Fundamentalvariablen gefunden wurden. Ein etwas anderes Bild ergibt sich beim Schätzen kurzfristiger Modelle. Sowohl die nichtlinearen neuronal entwickelten Modelle als auch lineare Strukturmodelle prognostizieren die kurzfristigen Veränderungen besser als ein Random Walk. Das zunächst motivierende Ergebnis wird durch ein zusätzliches Benchmarkmodell eines AR(6) Prozesses gedämpft. Die Prognosegüte des univariaten Zeitreihenmodells bewegte sich auf ähnlichem Niveau der nichtlinearen und der Strukturmodelle, so daß sich der zunächst vermutete positive Einfluß der Fundamentalvariablen relativiert.

5.5 GP GAUSS

Die Anwendung Genetischer Programme zur Entwicklung von Inflationsprognosemodellen haben gezeigt, daß für eine fixe In-Sample- und Out-of-Sample Datenmenge und verschiedenen Auswahlkriterien für die Prognosegleichungen die Ergebnisse recht ernüchternd waren. Selbst wenn eine der verschiedenen Auswahlstrategien zufällig auf eine vermeintlich gute Prognosegleichung gestoßen ist, so waren die Vorhersagen nicht signifikant. Dieses Ergebnis gilt sowohl für die Punkt- als auch für die Richtungsprognosen. Die zunächst besseren Prognosen beim Verwenden eines Informationskriteriums haben sich bei einem zweiten Lauf mit GP GAUSS nicht eingestellt. Darüber hinaus haben die bisherigen Untersuchungen gezeigt, daß für die Güte der Prognosegleichungen die Suchintensität nicht entscheidend ist und sogar kontraproduktiv wirken kann. Dies belegen Histogramme mit Prognoseverteilungen über die ersten 200 Generationen im Vergleich zu genetisch entwickelten Schätzgleichungen zwischen der 201. und 1000. Generation sowie die Analyse der Ergebnisse des Backpropagation Moduls. Insgesamt zeigen die Untersuchungen der Prognosegleichungen auf der Grundlage einer fixen In- und Out-of-Sample Datenmenge unbefriedigende Ergebnisse. Der bei allen Histogrammen zu beobachtende große prozentuale Anteil sehr schlechter Prognoseergebnisse signalisiert die Schwierigkeiten des Genetischen Programms bei der Extrapolation von Vorhersagen. Daher soll dieses Kapitel überprüfen, ob eine rekursive Regressionsmethode das Problem beheben kann. Die beiden Fitneßkriterien, MSE und Informationkriterium, sollen in diesem Zusammenhang erneut miteinander verglichen werden. Tabelle 5.2 enthält die Parametereinstellungen für GP GAUSS, die für die verschiedenen Simulationsläufe zugrundegelegt werden. Dabei wird die Anzahl der Generationen auf 200 und die Populationsgröße auf 100 Gleichungen beschränkt. GP GAUSS schätzt für Deutschland, Niederlande, Großbritannien und Japan die verschiedenen Wechselkursmodelle einmal für den gesamten Datensatz, der sich von Januar 1974 bis Dezember 1998 (Deutschland und Niederlande) bzw. Oktober 2003 (Japan

Anzahl der Läufe	=	abhängig vom Datensatz
Anzahl der Generationen pro Lauf	=	200
Populationsgröße	=	100
Fitneßkriterium	=	MSE oder AIC
maximale Länge der Funktionen der Anfangspopulation	=	5 Knoten
maximale Länge der Funktionen im weiteren Verlauf	=	100 Knoten
Anzahl neuer Gleichungen durch Crossoveroperationen	=	60
Anzahl neuer Gleichungen durch Mutationsoperationen	=	40
Typ A:	=	20
Typ B:	=	20
maximale Länge der Funktionsteile bei Mutationstyp B	=	5 Knoten
Anzahl der besten Funktionen, die für genetische Operationen herangezogen werden	=	100

Tabelle 5.2: Aufgrund rekursiver Schätzungen und dem Vorliegen von Datensätze mit unterschiedlichen Erhebungszeiträumen ist die Anzahl der Läufe pro Simulation abhängig vom Datensatz.

und Großbritannien) erstreckt, und anhand eines zweiten Erhebungszeitraumes beginnend ab März 1987 bis Dezember 1998 bzw. Oktober 2003. Die In-Sample Menge für den Datensatz beginnend ab 1974 startet mit 180 Werten bis einschließlich Dezember 1988. Für den zweiten Datensatz beginnend ab 1987 werden die ersten 82 Werte bis einschließlich Dezember 1993 zum Schätzen des ersten Modells verwendet. Diese Einteilung gilt für alle untersuchten Zeitreihen. Aufgrund der längeren Zeitreihen bei Großbritannien und Japan werden für diese beiden Staaten im Vergleich zu Deutschland und den Niederlanden insgesamt 58 zusätzliche Prognosen erstellt. Das Genetische Computerprogramm endet, sobald die In-Sample Menge der Out-of-Sample Menge abzüglich des kleinsten Prognosehorizonts entspricht. Daher liegen für die Berechnung des RMSE für die verschiedenen Prognosehorizonte auch eine unterschiedliche Anzahl von Prognosewerten zugrunde. Durch das Verwenden der Daten ab März 1987 entfallen einerseits einschneidende Ereignisse wie die beiden Ölpreiskrisen aus den Jahren 1973 und 1979 sowie die massiven Notenbankinterventionen im Zeitraum September 1985 bis Februar 1987, die das Auffinden guter Prognosegleichungen stark erschweren könnten. Es soll überprüft werden, ob die Phase ab März 1987 die Ergebnisse qualitativ verbessert. Diese beiden Datensätze werden mit dem MSE und dem AIC Kriterium für die kurz- und langfristigen Modelle geschätzt, so daß sich insgesamt 96 Modelle ergeben⁹. Zum Schätzen der Strukturmodelle anhand des Genetischen Programms gehen die ökonomischen Fundamentalvariablen mit bis zu sechs Zeitverzögerungen und die Wechselkurzeitreihe mit ebenfalls 6 Lags in die Terminalmenge ein, wie sie in der Gleichung 5.20 aufgeführt sind.

⁹ 4 Staaten * 2 Fitneßkriterien * 2 Datensätze * lang- und kurzfristige Modelle * 3 Wechselkursmodelle = 96 genetisch zu schätzende Modelle

Modell	1 Monat	6 Monate	12 Monate	18 Monate	24 Monate
Japan					
Flexible-Price	0.030	0.094	0.129	0.179	0.228
Sticky-Price	0.029	0.090	0.115	0.157	0.272
Portfolio	0.034	0.099	0.124	0.164	0.192
Random Walk	0.028	0.085	0.111	0.145	0.169
Großbritannien					
Flexible-Price	0.024	0.068	0.089	0.099	0.117
Sticky-Price	0.023	0.066	0.090	0.130	2.960
Portfolio	0.024	0.066	0.086	0.089	0.104
Random Walk	0.023	0.066	0.086	0.090	0.102
Deutschland					
Flexible-Price	0.0262	0.0767*	0.107*	0.116*	0.120*
Sticky-Price	0.0266	0.0821	0.115	0.133*	0.186*
Portfolio	0.0303	0.0807	0.115	0.537	3485.872
Random Walk	0.0261	0.0775	0.105	0.119	0.123
Niederlande					
Flexible-Price	0.026	0.078	0.307	0.343	0.369
Sticky-Price	0.027	0.078	0.106	0.232	0.270
Portfolio	0.027	0.085	0.236	0.249	0.264*
Random Walk	0.026	0.077	0.106	0.121	0.125

Tabelle 5.3: langfristige Prognosemodelle; RMSE Punktprognose mit **AIC als Fitneßkriterium**; In-Sample Zeitraum: Januar 1974 bis Dezember 1988. Out-of-Sample Zeitraum: Januar 1989 bis Dezember 1998 bzw. Oktober 2003. *Ein ∞ Wert (**zwei Werte) (***)drei Werte) wurde durch einen naiven Wert ersetzt.

5.6 Ergebnisse der lang- und kurzfristigen Prognosemodelle

Die umfangreichen Ergebnisse beginnen mit der Analyse des ersten Datensatzes mit Beobachtungswerten zwischen Januar 1974 bis Dezember 1998 für Deutschland und den Niederlanden bzw. Oktober 2003 für Japan und Großbritannien. Die anhand eines MSE und AIC-Fitneßkriteriums genetisch entwickelten Punkt- und Richtungsprognosen sind in den Tabellen 5.3 und 5.4 respektive 5.5 und 5.6 für die Prognosehorizonte $k = 1, 6, 12, 18, 24$ Monate abgebildet¹⁰. Vergleicht man die Punktprognosen mit den Ergebnissen von Verkooijen et al.¹¹, die die Autoren mit Hilfe künstlicher neuronaler Netze für die Periode bis Juli 1994 erhalten haben, so fällt auf, daß sich die Ergebnisse qualitativ nicht voneinander unterscheiden. Unabhängig von der Wahl des Fitneßkriteriums ist

¹⁰Alle in diesem Kapitel angegebenen Wechselkursvorhersagen Genetischer Programme wurden anhand der besten In-Sample Regressionsgleichung erzeugt.

¹¹Die Autoren haben anhand eines künstlichen neuronalen Netzwerkes für die gleichen Länder und Prognosehorizonte die Wechselkursmodelle rekursiv geschätzt. Keines ihrer langfristigen, nichtlinearen Modelle konnte bessere Prognosen als ein Random Walk erzielen. Siehe auch Anhang D.8.

Modell	1 Monat	6 Monate	12 Monate	18 Monate	24 Monate
Japan					
Flexible-Price	8.854	9.003	9.182	9.367	9.570
Sticky-Price	0.029	0.105	0.148	0.541	0.576*
Portfolio	0.030	0.095*	0.139*	0.187*	0.222*
Random Walk	0.028	0.085	0.111	0.145	0.169
Großbritannien					
Flexible-Price	0.024	1.009**	1.033***	1.053***	1.077****
Sticky-Price	11.65	11.88*	12.11**	12.34**	12.55**
Portfolio	0.025	0.069	0.093	0.102	0.130
Random Walk	0.023	0.066	0.086	0.090	0.102
Deutschland					
Flexible-Price	0.0264*	0.086*	0.122*	0.175**	0.199**
Sticky-Price	0.0268	0.123	0.287**	0.863**	1.668***
Portfolio	0.0268*	0.085*	0.122*	0.151*	0.208*
Random Walk	0.0261	0.0775	0.105	0.119	0.123
Niederlande					
Flexible-Price	0.025	0.082*	0.125*	0.173*	0.220
Sticky-Price	0.027	0.109*	0.152*	0.193*	0.410*
Portfolio	0.025	0.085	0.117	25.64	26.49
Random Walk	0.026	0.077	0.106	0.121	0.125

Tabelle 5.4: langfristige Prognosemodelle; RMSE Punktprognose mit **MSE als Fitneßkriterium**; In-Sample Zeitraum: Januar 1974 bis Dezember 1988. Out-of-Sample Zeitraum: Januar 1989 bis Dezember 1998 bzw. Oktober 2003. *Ein ∞ Wert (**zwei Werte) (***)drei Werte) wurde durch einen naiven Wert ersetzt.

kein genetisch entwickeltes Strukturmodell in der Lage, die Vorhersagefehler eines naiven Random Walk Modells zu unterbieten. Ungeachtet der schlechten Vorhersagen hebt sich auch beim internen Vergleich der Strukturmodelle kein Ansatz positiv hervor. Die bei der Untersuchung von Inflationsprognosen beobachteten unrealistischen Prognosen genetisch entwickelter Modelle sind bei den Wechselkursprognosen ebenfalls anzutreffen. Um einigermaßen vernünftige Werte zu erhalten, wurden daher unendliche Wechselkursvorhersagen durch den Vorhersagewert eines Random Walks ersetzt. Die davon betroffenen Prognosewerte sind mit einem '*' gekennzeichnet, wobei die Anzahl der Symbole zugleich die Menge an ausgetauschten Werten beschreibt. In diesem Zusammenhang weist das Informationskriterium eine deutlich geringere Anzahl solcher Prognosewerte auf als das übliche MSE Kriterium. Letzteres mußte insgesamt 50mal korrigiert werden im Gegensatz zu den 7 substituierten Prognosewerten beim Informationskriterium, wobei die Zeitreihen aus Deutschland und Großbritannien die mit Abstand größten Probleme bereiteten. Vergleicht man die nicht durch Ausreißerwerte verzerrten Ergebnisse des MSE und des Informationskriteriums miteinander, dann ähneln sich deren Vorhersagen bei kleinen Prognosehorizonten bis 6 Monaten. Dies ändert sich tendenziell bei Prognosen über 18 und 24 Monaten

zugunsten des Informationskriteriums. Im Anhang auf den Seiten 172 bis 175 sind die Wechselkursprognosen der beiden Fitneßkriterien für den 24monatigen Prognosehorizont abgebildet. Aus den Graphiken ist ersichtlich, daß GP GAUSS - trotz teilweise ersetzter Prognosewerte - extreme Ausreißerwerte generiert. Insgesamt zeigen die Ergebnisse, daß das Informationskriterium gegenüber dem MSE als Fitneßkriterium Vorteile aufweist. Im Vergleich zu künstlichen neuronalen Netzwerken muß festgestellt werden, daß aufgrund der Problematik extremer Wechselkursprognosen bei GP GAUSS der Suchalgorithmus künstlicher neuronaler Netzwerke zumindest für die vorliegende Untersuchung vorzuziehen ist. Die zusätzlich zu den Punktprognosen berechneten Richtungsprognosen liefern etwas erfreulichere Ergebnisse und sind in den Tabellen 5.5 und 5.6 dargestellt. Da bei den Richtungsprognosen lediglich die Richtung und nicht das Ausmaß der Änderungen vorhergesagt werden muß, fallen die Anforderungen an die Prognosemodelle geringer aus als bei den Punktprognosen. Beim Blick auf die beiden Tabellen stechen die sehr guten und hoch signifikanten einmonatigen Richtungsprognosen hervor. Demnach sagen die genetischen Modelle unabhängig vom gewählten Fitneßkriterium die Richtungsänderungen für den nächsten Monat mit Werten zwischen 74% und 86% korrekt voraus und übertreffen die Prognosen des Random Walks, dessen Werte zwischen 52% und 62% betragen, deutlich. Mit zunehmenden Prognosehorizont verringert sich zwar die prozentuale Vorhersagegüte einhergehend mit erheblich niedrigeren Signifikanzniveaus, aber sie liegen nach wie vor mit wenigen Ausnahmen sichtbar über den Werten eines naiven Modells. Auffallend sind ebenfalls die zum Teil sehr signifikanten Richtungsprognosen für den 12monatigen Prognosehorizont. Das Verwenden eines MSE- oder eines Informationskriteriums als Fitneßkriterium verursacht bei den Richtungsprognosen keine erkennbaren Unterschiede. Bei einem erneuten Vergleich der Strukturmodelle untereinander lassen sich nach wie vor keine Präferenzen zugunsten eines theoretischen Ansatzes feststellen.

Die beiden Tabellen 5.7 und 5.8 enthalten die Ergebnisse der kurzfristigen Modelle, die auf der Grundlage differenzierter Zeitreihenwerte geschätzt worden sind. Die Ein-Monats-Punktprognosen für die Änderungsraten des Wechselkurses stimmen mit den Ergebnissen von Plasman et al. qualitativ überein. Alle Genetischen Programme erzeugen bessere Prognosen als ein Random Walk. Allerdings zeigt der modifizierte Diebold-Mariano Test, daß die GP GAUSS Wechselkursprognosen sich nicht signifikant von den Random Walk Vorhersagen unterscheiden. Im Gegensatz dazu sind die Richtungsänderungen der Wechselkursänderungsraten erneut hoch signifikant und ähneln vom Ausmaß her den langfristigen Modellen bei der Ein-Monats-Richtungsprognose.

Die vorangegangenen Ergebnisse basieren auf Zeitreihen, die die Finanzmarkturbulenzen seit dem Zusammenbruch des Bretton-Woods Systems sowie die verschiedenen zentralbankpolitischen Eingriffe widerspiegeln. In einer weiteren Untersuchung soll daher überprüft werden, ob sich die Resultate bei Verwenden von Zeitreihen nach Beendigung der massiven Zentralbankinterventionen zu Beginn der 1980er Jahre verbessern. Der Zeitraum von März 1987 bis Dezember 1998 bzw. Oktober 2003 stellt eine ruhigere Periode dar sowohl hinsichtlich exogener Schocks als auch in Bezug auf das Ausmaß der Zentralbankinterventionen auf den Devisenmärkten. Die Darstellung der Ergebnisse beginnt wieder mit der Betrachtung der Punktprognosen, die für die beiden Fitneßkriterien jeweils in den

Tabellen D.2 und D.4 im Anhang aufgeführt sind¹². Wie schon beim ersten Datensatz fällt auf, daß beim MSE Kriterium insgesamt 38 schlechte Prognosewerte durch Random Walk Vorhersagen substituiert werden müssen im Vergleich zu 9 Werten beim Informationskriterium, so daß das Informationskriterium wiederum verlässlichere Vorhersagen hervorruft als das MSE Kriterium. Dies macht sich auch erneut bei den langfristigen Prognosewerten bemerkbar, die eine tendenziell genauere Vorhersagegüte erkennen lassen. Allerdings gelingt es den Modellen aufs Neue nicht, bessere Vorhersagen als ein naives Modell zu erzeugen. Etwas besser sehen die Richtungsprognosen in den Tabellen D.3 und D.5 (Seite 167 und 169) aus, die leicht höhere prozentuale Werte für korrekte Richtungsänderungen mit zum Teil verbesserten Teststatistiken wiedergeben. Trotzdem weichen die Ergebnisse insgesamt betrachtet nicht wesentlich von den anhand des ersten Datensatzes erhaltenen Vorhersagewerten ab.

Zum Schluß richtet sich der Blick erneut auf die Ergebnisse der kurzfristigen Modelle, die in den beiden Tabellen D.6 und D.7 (Seite 170 und 171) abgebildet sind. Auffällig sind die negativen Abweichungen der Vorhersagegüte des Portfolioansatzes, die sich ohne Ausnahme bei allen Staaten sowohl beim MSE als auch beim Informationskriterium einstellen. Während die Richtungsprognosen trotz einer schlechteren Trefferquote gegenüber den beiden monetären Ansätzen zumindest noch signifikant über den Werten des Random Walks liegen, fallen die Vorhersagefehler der Punktprognose deutlich schlechter aus.

5.7 Zusammenfassung

In diesem Kapitel wurden Wechselkursprognosen auf der Grundlage monetärer und portfoliotheoretischer Bestimmungsgleichungen aufgestellt, wobei sich die Vorgehensweise an den Artikeln von Plasmans et al. [104] und Verkoojiens et al. [127] anlehnt hat. Während sich die Autoren auf Punktprognosen eines ausgewählten Zeitraumes beschränken, wurden die Untersuchungen auf Richtungsprognosen und dem Betrachten zweier unterschiedlicher Zeitperioden ausgedehnt. Der erste Zeitraum erstreckt sich von Januar 1974 bis Dezember 1998 bzw. Oktober 2003 und umfaßt einschneidende Ereignisse wie die Ölkrisen in den 1970er Jahren oder die auf das Plaza Agreement zurückgehenden koordinierten Zentralbankinterventionen Mitte der 1980er Jahre. Die Daten des zweiten Zeitraumes setzen dagegen im Anschluß an die Interventionsphase ab März 1987 an, so daß eine Zeitperiode ohne vergleichbar große exogenen Schocks bzw. zentralbankpolitische Eingriffe auf den Finanzmärkten zugrundeliegt. Die Untersuchungen dieses Kapitels verfolgten mehrere Ziele. Die Anlehnung der Untersuchungen an die Aufsätze von Plasmans et al. und Verkoojiens et al. ermöglichten einen Vergleich zwischen künstlichen neuronalen Netzen und Genetischen Programmen. Es zeigt sich, daß das Verwenden unterschiedlicher Suchalgorithmen auf die gleiche Problemstellung Ergebnisse mit qualitativ gleichwertigen Aussagen erzeugt. Dies ist insofern interessant, da sich die beiden nichtlinearen Methoden sowohl in ihrem Verfahren als auch im Hinblick auf die ermittelten Schätzgleichungen grundsätzlich voneinander unterscheiden, wie bereits im ersten Kapitel der Arbeit dargelegt worden ist. Die Ergebnisse belegen sowohl bei Plasmans et al. als auch bei genetisch entwickelten

¹²Siehe Anhang D, Seite 166!

Modellen, daß langfristige Strukturmodelle bei Punktprognosen dem naiven Modell eines Random Walks unterlegen sind und nur bei kurzfristigen Prognosemodellen eine schwache Evidenz für die theoretisch postulierten Zusammenhänge ökonomischer Fundamentalvariablen unterbreiten. Allerdings sind die Ergebnisse nicht signifikant. Die letzte Vermutung wird allerdings durch die Verwendung des zweiten Datensatzes 1987 bis 1998 bzw. 2003 etwas erschüttert, wonach der portfoliotheoretische Ansatz bei der Vorhersage einmonatiger Wechselkursänderungen bei allen vier Staaten und unabhängig vom gewählten Fitneßkriterium größere Fehlerabweichungen als ein naives Modell erzeugt¹³. Obwohl die Aussagen sich qualitativ nicht unterscheiden, erweist sich das künstliche neuronale Netz gegenüber dem Genetischen Programmieren vor allem aus einem Grunde als vorteilhaft. Die Genetischen Programme erzeugten relativ häufig unendlich große und somit unbrauchbare Prognosewerte, die zwar durch das Verwenden eines Informationskriteriums gemildert aber nicht beseitigt werden konnten. Als positives Ergebnis dieser Untersuchungen stehen die sehr guten genetisch entwickelten Richtungsprognosen hervor. Vor allem für einmonatige Richtungsvorhersagen erzeugte GP GAUSS mit Werten zwischen 74% und 86% stark signifikante Prognosen, die sich deutlich gegenüber den Vorhersagen eines Random Walks (52% und 62%) abheben.

¹³Vgl. Anhang D, S. 170.

Modell	1 Monat	6 Monate	12 Monate	18 Monate	24 Monate
Japan					
Flexible-Price	81.66 (33.87)	54.27 (0.59)	62.66 (5.06)	57.89 (1.89)	45.89 (0.49)
Sticky-Price	84.02 (39.14)	57.93 (2.07)	62.03 (4.59)	52.63 (0.22)	54.11* (0.48)
Portfolio	81.07 (32.79)	51.22* (0.049)	58.86* (2.48)	57.89* (1.85)	58.90* (2.30)
Random Walk	56.80 (1.53)	48.17 (0.12)	45.57 (0.63)	50.66 (0.01)	43.53 (1.24)
Großbritannien					
Flexible-Price	86.39 (44.91)	59.76** (3.17)	68.35*** (10.65)	63.82*** (5.88)	57.53**** (1.68)
Sticky-Price	84.62 (40.85)	66.46* (8.94)	63.29** (5.69)	55.92** (1.07)	61.64** (4.06)
Portfolio	85.80 (43.35)	61.59 (4.43)	62.03 (4.54)	62.50 (4.75)	57.53 (1.63)
Random Walk	52.07 (0.13)	50.0 (0.00)	50.0 (0.00)	53.95 (0.47)	52.74 (0.21)
Deutschland					
Flexible-Price	80.18* (20.25)	56.60* (0.94)	58.00* (1.28)	61.70** (2.52)	64.77** (3.83)
Sticky-Price	79.28 (19.35)	55.67 (0.66)	47.00** (0.17)	54.23** (0.30)	50.00*** (0.00)
Portfolio	80.18* (20.25)	60.38* (2.28)	63.00* (3.34)	52.13* (0.09)	56.82* (0.82)
Random Walk	62.16 (3.28)	53.77 (0.30)	50.00 (0.00)	37.23 (3.08)	41.57 (1.26)
Niederlande					
Flexible-Price	81.08 (21.63)	60.38* (2.26)	61.00* (2.44)	53.20* (0.19)	52.27 (0.08)
Sticky-Price	78.37 (18.11)	56.60* (0.90)	64.00* (3.93)	51.06* (0.02)	60.22* (1.80)
Portfolio	76.58 (15.89)	65.09 (4.82)	57.00 (0.98)	54.26 (0.34)	51.14 (0.02)
Random Walk	60.36 (2.32)	53.77 (0.29)	52.00 (0.08)	39.36 (2.14)	46.07 (0.32)

Tabelle 5.5: langfristige Prognosemodelle; Richtungsprognosen in Prozent; **MSE als Fit-
neßkriterium**; In-Sample Zeitraum: Januar 1974 bis Dezember 1988. Out-of-Sample Zeit-
raum: Januar 1989 bis Dezember 1998 bzw. Oktober 2003. Signifikanzniveaus von 10%,
5% und 1% entsprechen Teststatistiken (in Klammern) von 2.71, 3.84 und 6.63.

Modell	1 Monat	6 Monate	12 Monate	18 Monate	24 Monate
Japan					
Flexible-Price	83.43 (37.86)	57.32 (1.73)	60.76 (3.65)	55.26 (0.82)	54.79 (0.68)
Sticky-Price	83.34 (37.86)	50.00 (0.00)	64.56 (6.73)	60.53 (3.29)	49.32 (0.01)
Portfolio	83.43 (37.97)	57.93 (2.03)	60.76 (3.66)	61.18 (3.77)	62.33 (4.45)
Random Walk	56.80 (1.53)	48.17 (0.12)	45.57 (62.83)	50.66 (0.01)	43.84 (1.13)
Großbritannien					
Flexible-Price	83.43 (37.86)	65.85 (8.32)	66.46 (8.55)	56.58 (1.35)	53.42 (0.37)
Sticky-Price	74.56 (20.46)	45.12 (0.74)	48.10 (0.12)	45.39 (0.64)	48.63 (0.06)
Portfolio	85.21 (41.96)	67.07 (9.58)	68.35 (10.63)	59.87 (3.01)	59.59 (2.75)
Random Walk	52.07 (0.13)	50.0 (0.00)	50.0 (0.00)	53.95 (0.47)	52.38 (0.16)
Deutschland					
Flexible-Price	78.38 (17.90)	66.04* (5.47)	64.00* (3.93)	57.45* (1.15)	51.14* (0.03)
Sticky-Price	78.38 (17.88)	61.32 (2.76)	62.00 (2.92)	47.87* (0.09)	51.14* (0.02)
Portfolio	77.48 (16.81)	65.09 (4.28)	68.00 (6.49)	61.70 (2.52)	63.64 (3.27)
Random Walk	62.16 (3.28)	53.77 (0.30)	50.00 (0.00)	37.23 (3.08)	41.57 (1.26)
Niederlande					
Flexible-Price	80.18 (20.48)	57.55 (1.22)	62.00 (2.95)	47.87 (0.09)	52.27 (0.09)
Sticky-Price	80.18 (20.33)	59.43 (1.86)	60.00 (2.03)	61.70 (2.59)	60.23 (1.83)
Portfolio	81.98 (22.98)	59.43 (1.90)	63.00 (3.38)	56.38 (0.77)	57.96* (1.16)
Random Walk	60.36 (2.32)	53.77 (0.29)	52.00 (0.08)	39.36 (2.14)	46.07 (0.32)

Tabelle 5.6: langfristige Prognosemodelle; Richtungsprognosen in Prozent; **AIC als Fitneßkriterium**; In-Sample Zeitraum: Januar 1974 bis Dezember 1988. Out-of-Sample Zeitraum: Januar 1989 bis Dezember 1998 bzw. Oktober 2003. Signifikanzniveaus von 10%, 5% und 1% entsprechen Teststatistiken (in Klammern) von 2.71, 3.84 und 6.63.

Modell	JP	GB	BRD	NL
AIC ab Januar 1974				
Flexible-Price	88.10 (48.86)	80.95 (32.18)	79.09 (18.68)	77.27 (16.51)
Sticky-Price	86.90 (45.79)	79.76 (29.71)	79.09 (18.68)	79.09 (18.95)
Portfolio	88.10 (48.84)	80.95 (32.16)	78.18 (17.57)	76.36 (15.37)
Random Walk	30.36 (12.97)	41.67 (2.52)	46.36 (0.30)	48.18 (0.10)
MSE ab Januar 1974				
Flexible-Price	89.29 (51.96)	79.17 (28.59)	80.91 (21.08)	77.27 (16.51)
Sticky-Price	82.74 (36.02)	82.74 (36.05)	77.27 (16.51)	77.27 (16.51)
Portfolio	88.69* (50.36)	79.76 (29.74)	77.27 (16.65)	77.27 (16.51)
Random Walk	30.36 (12.97)	41.67 (2.52)	46.36 (0.30)	48.18 (0.10)

Tabelle 5.7: kurzfristige Prognosemodelle; Richtungsprognosen in Prozent; In-Sample Zeitraum: Januar 1974 bis Dezember 1988. Out-of-Sample Zeitraum: Januar 1989 bis Dezember 1998 bzw. Oktober 2003. *Ein ∞ Wert (**zwei Werte) (***)drei Werte) wurde durch einen naiven Wert ersetzt. Signifikanzniveaus von 10%, 5% und 1% entsprechen Teststatistiken (in Klammern) von 2.71, 3.84 und 6.63.

Modell	JP	GB	BRD	NL
AIC ab Januar 1974				
Flexible-Price	0.0283 (0.340)	0.0235 (0.360)	0.0251 (0.332)	0.0266 (0.375)
Sticky-Price	0.0289 (0.378)	0.0239 (0.373)	0.0249 (0.338)	0.0221 (0.34)
Portfolio	0.0271 (0.359)	0.0228 (0.328)	0.0265 (0.353)	0.0255 (0.34)
Random Walk	0.0330	0.0278	0.0298	0.0298
MSE ab Januar 1974				
Flexible-Price	0.0284 (0.328)	0.0232 (0.328)	0.0257 (0.348)	0.0259 (0.35)
Sticky-Price	0.0296 (0.415)	0.0221 (0.323)	0.0278 (0.350)	0.0254 (0.35)
Portfolio	0.0323* (0.395)	0.0232 (0.343)	0.0279 (0.343)	0.0262 (0.35)
Random Walk	0.0330	0.0278	0.0298	0.0298

Tabelle 5.8: kurzfristige Prognosemodelle; In-Sample Zeitraum: Januar 1974 bis Dezember 1988. Out-of-Sample Zeitraum: Januar 1989 bis Dezember 1998 bzw. Oktober 2003. Die mit GP GAUSS erzeugten RMSE Punktprognosen liefern zwar bessere Werte als ein Random Walk, wobei diese allerdings nicht signifikant vom naiven Modell abweichen (Signifikanzniveaus in Klammern). *Ein ∞ Wert (**zwei Werte) (***)drei Werte) wurde durch einen naiven Wert ersetzt.

Zusammenfassung

Die Anwendung des Genetischen Programmierens auf monetäre Größen diente dem Ziel, die Prognosefähigkeit dieser speziellen nichtlinearen Regressionsmethode zu überprüfen. Hierzu wurde auf der Grundlage des Matrizenrechenprogramms GAUSS ein entsprechendes Computerprogramm entwickelt.

Das mit GP GAUSS bezeichnete Programm wurde in einem ersten Schritt auf verschiedene nichtlineare dynamische Prozesse angewandt und mit den Ergebnissen anderer Genetischer Programme verglichen. In diesem Vergleich der In-Sample Schätzungen war GP GAUSS in der Lage, die nichtlinearen Prozesse zum Teil deutlich besser zu regressieren. Eine anschließende Sensitivitätsanalyse untersuchte verschiedene Faktoren, um das Genetische Programm zu optimieren. Die Ergebnisse aus dieser Untersuchung wurden bei den späteren Anwendungen auf monetäre Größen weitestgehend berücksichtigt. Innerhalb dieser Analyse wurden zwei besondere Untersuchungen vorgestellt. Die erste Untersuchung stellte ein Backpropagation Verfahren dar, mit dessen Hilfe die einzelnen Terme einer genetisch entwickelten Gleichung weiter optimiert werden konnten. Die zweite besondere Untersuchung dieses Kapitels war die Analyse der Building Blocks. Mit der hierfür entwickelten Prozedur innerhalb des Computerprogramms konnte der Einfluß eines einzelnen Building Blocks auf die Fitneßänderung exakt nachgewiesen werden. Demnach nimmt die Fitneßänderung tendenziell mit der Zunahme der Länge eines ausgetauschten Teilbereichs ab. Zudem konnten die zeitlich nachgelagerten Auswirkungen der Building Blocks nach einer größeren sprunghaften In-Sample Verbesserung graphisch veranschaulicht werden.

Ausgestattet mit diesen Ergebnissen prognostizierte GP GAUSS im nächsten Kapitel auf Basis des Mishkin Ansatzes deutsche Inflationsraten. Dabei wurden die Schätzungen und Prognosen anhand eines fixen In- und Out-of-Sample Datensatzes vorgenommen. Mit den Ergebnissen der In-Sample Schätzungen wurden verschiedene Strategien zur Auswahl der Prognosegleichung für die Out-of-Sample Vorhersage untersucht. Ein zunächst interessantes Ergebnis lieferte die Anwendung eines Informationskriteriums als Fitneßkriterium, das neben der Fehlerabweichung zusätzlich die Länge einer Prognosegleichung in der Bewertung berücksichtigt. Die auf Grundlage dieses Informationskriteriums in einem Histogramm dargestellte Verteilung der Prognosegüte genetisch entwickelter Gleichungen war besser als die Prognoseverteilung bei Verwenden einer mittleren quadratischen Fehlerabweichung als Fitneßkriterium. Allerdings konnte eine wiederholte Anwendung von GP GAUSS diese Ergebnisse nicht bestätigen. Darüber hinaus veranschaulichten die Histogramme, daß GP GAUSS überhaupt Regressionsgleichungen entwickeln konnte, die bessere Punktprognosen als ein Random Walk und andere Vergleichsmodelle erzeugen kann.

Die Wahl einer guten Prognosegleichung aus der Menge an potentiellen Lösungskandidaten mit Hilfe einer bestimmten Strategie war allerdings zufallsbedingt. Der modifizierte Diebold-Mariano Test signalisierte zudem, daß keine der untersuchten, besseren Punktprognosen sich signifikant von den Vorhersagewerten eines Random Walks unterscheiden haben. Das Gleiche gilt für die Richtungsprognosen.

Im letzten Kapitel der Arbeit wird das Genetische Programm auf monetaristische und portfoliobasierte Wechselkursmodelle angewendet, wobei die Wechselkursprognosen im Unterschied zu den Inflationsprognosen rekursiv ermittelt werden. Das Kapitel orientiert sich dabei stark an die Artikel von Plasmans et al. [104] und Verkooijens [127] et al., die die Wechselkurse mit Hilfe künstlicher neuronaler Netzwerke geschätzt haben. Es sollte überprüft werden, ob unterschiedliche nichtlineare Schätzverfahren zu den gleichen qualitativen Aussagen gelangen. Erneut wurde das Informationskriterium neben der mittleren quadratischen Fehlerabweichung als Fitneßkriterium herangezogen. Es zeigt sich, daß für langfristige Prognosehorizonte der Random Walk nicht und für kurzfristige Prognosehorizonte das naive Modell kaum bessere Punktprognosen erzeugt. Damit folgen die Ergebnisse des Genetischen Programms denen von Plasmans et al.. Allerdings zeigt der modifizierte Diebold Mariano Test für die kurzfristigen Vorhersagen erneut, daß diese nicht signifikant von den Vorhersagen eines Random Walks abweichen. Dagegen sind die Richtungsprognosen für die langfristigen und kurzfristigen Prognosemodelle für den Ein-Monats-Prognosehorizont stark signifikant. Es muß jedoch betont werden, daß für die in der Regressionsgleichung benötigten zukünftigen Variablen die aktuellen Werte (wie bei Plasmans et al. und Verkooijens et al.) verwendet wurden, so daß auch dieses auf den ersten Blick positive Ergebnis relativiert wird. Darüber hinaus mußten für die Prognoseberechnungen einige unendlich große Prognosewerte durch die Werte eines Random Walks ersetzt werden, um vernünftigeren Vorhersagewerte zu bekommen. Bei Verwenden des Informationskriteriums konnte die Anzahl der zu ersetzenden Werte reduziert werden.

Insgesamt kann festgehalten werden, daß das Genetische Programm gute In-Sample Schätzungen erzeugt. Bei den Out-of-Sample Vorhersagen liefern zumindest die rekursiv ermittelten Wechselkurs-Richtungsprognosen ermutigende Ergebnisse. Weitergehende Untersuchungen mit neuen Ideen könnten die Anwendung Genetischer Programme auf monetäre Größen eventuell weiter verbessern.

Anhang A

Tests auf Nichtlinearität

Abgesehen vom Bisketraltest werden die Zeitreihen mit Hilfe eines AR(p)-Prozesses zunächst von möglichen linearen Abhängigkeiten bereinigt, wobei die Laglänge p auf der Grundlage des Schwarz Informationskriteriums bestimmt wird. Die durch diesen Prewhitening Vorgang erhaltenen Schätzresiduen bilden den Ausgangspunkt für die weitere Vorgehensweise der verschiedenen nichtlinearen Testverfahren. Dabei wird stets die Nullhypothese unabhängig und identisch verteilter Fehlerterme überprüft.

A.1 McLeod & Li

Das Testverfahren¹ von McLeod und Li (1983) [88] überprüft, ob in den quadrierten Fehlertermen u_t^2 des Prewhitening Prozesses mögliche ARCH Effekte enthalten sind. Hierzu werden für verschiedene Zeitverzögerungen k die Autokorrelationswerte $\rho(k)$ berechnet:

$$\rho(k) = \frac{\sum_{k=1}^L (u_t^2 - \sigma^2)(u_{t-k}^2 - \sigma^2)}{\sum_{t=1}^n (u_t^2 - \sigma^2)} \quad (\text{A.1})$$

mit $\sigma^2 = \sum_{t=1}^N \frac{u_t^2}{N}$. Anhand der Box-Ljung Teststatistik wird die Nullhypothese untersucht, ob die Autokorrelationswerte für verschiedene Lags k von Null verschieden sind:

$$Q = n(n+2) \sum_{i=1}^L \frac{\rho(k)^2}{n-i} \quad (\text{A.2})$$

Die Teststatistik folgt dabei einer $\chi^2(L)$ Verteilung. Die weiter unten aufgeführten Ergebnisse wurden für $L = 24$ berechnet.

A.2 Engle

Engle [33] (1982) entwickelte im Zusammenhang mit den von ihm aufgestellten ARCH Modellen ein Testverfahren, das die Fehlerterme auf mögliche ARCH Effekte untersucht.

¹Die Erläuterungen der nachfolgenden Testverfahren orientieren sich größtenteils an Patterson & Ashley [3], S. 39 ff.

Es handelt sich daher um einen Test auf Nichtlinearität in den Varianzen und nicht um die Bestimmung allgemeiner Nichtlinearität. Zum besseren Verständnis soll ein ARCH(p) in seiner einfachsten Form dargestellt werden:

$$y_t = c + \phi_1 y_{t-1} + \phi_2 y_{t-2} + \dots + \phi_m y_{t-m} + v_t \quad (\text{A.3})$$

$$v_t \sim N(0, \sigma_t) \quad (\text{A.4})$$

$$\sigma_t^2 = \alpha_0 + \alpha_1 v_{t-1}^2 + \alpha_2 v_{t-2}^2 + \dots + \alpha_p v_{t-p}^2 \quad (\text{A.5})$$

Bezogen auf die - durch den Prewhitening Prozeß erhaltenen - Schätzresiduen u_t soll getestet werden, ob die Fehlerterme ARCH Effekte aufweisen. Hierzu werden die quadrierten Schätzresiduen auf eine Konstante und zeitverzögerte Variablen von u_t^2 regressiert:

$$u_t^2 = \alpha_0 + \sum_{i=1}^k \alpha_i u_{t-i}^2 + \epsilon_t \quad (\text{A.6})$$

Wird die Nullhypothese keiner Autokorrelation $H_0 : \alpha_1 = \alpha_2 = \dots \alpha_k = 0$ unterstützt, dann ist $n * R^2$ mit R^2 als Bestimmtheitsmaß asymptotisch $\chi^2(k)$ verteilt. Die in der Tabelle aufgeführten Werte sind für Zeitverzögerungen bis $p = 5$ berechnet worden.

A.3 Tsay

Beim Test von Tsay (1986) [125] stehen die Kumulanten 3. Ordnung der stationären Zeitreihe x_t im Mittelpunkt der Betrachtung². Hierfür wird zunächst ein Vektor z generiert, der die Elemente der Diagonalmatrix von $[x_{t-1} \dots x_{t-M}]' [x_{t-1} \dots x_{t-M}]$ enthält³. Anschließend wird der Vektor z_t auf den Vektor $w_t = [1 \ x_{t-1} \dots x_{t-M}]$ regressiert und die Fehlerabweichungen a_t werden gespeichert. Schließlich erfolgt die Regression der Fehlerabweichungen u_t (aus dem Prewhitening Prozeß) auf a_t :

$$u_t = \gamma_0 + \sum_{i=1}^M \gamma_i a_{ti} + \eta_t \quad (\text{A.7})$$

Mit Hilfe einer F-Statistik überprüft die Tsay Teststatistik die Nullhypothese, ob die Koeffizienten γ_1 bis γ_M den Wert null annehmen.

A.4 Hinich Bikovarianz Test

Unter der Annahme, daß u_t einem stationären stochastischen Prozeß dritter Ordnung entspricht, werden beim Test von Hinich (1996) [63] die Bikovarianzen dieser Zeitreihe auf Unkorreliertheit überprüft. Die Bikovarianz, die mit dem Schiefe-Parameter einer Verteilung übereinstimmt, ist definiert als:

$$C_3(r, s) = (N - s)^{-1} \sum_{t=1}^{N-s} u_t u_{t+r} u_{t+s} \quad (\text{A.8})$$

²Siehe auch Brock et al. [15], S.59 und Ashley et al. [3], S.41 f.

³Bsp. für $M = 2$: $z_t = [x_{t-1}^2 \ x_{t-1} x_{t-2} \ x_{t-2}^2]$.

mit $0 \leq r \leq s$. Es wird die Nullhypothese einer identisch und unabhängig verteilten Zeitreihe u_t überprüft ($C_3(r, s) = 0$). Für $G(r, s) = (N - s)^{1/2} C_3(r, s)$ definiere X_3

$$X_3 = \sum_{s=2}^l \sum_{r=1}^{s-1} [G(r, s)]^2 \quad (\text{A.9})$$

Unter der Nullhypothese eines i.i.d. verteilten Prozesses für x_t zeigen Hinich & Patterson (1995), daß X_3 asymptotisch $\chi^2(l[l-1]/2)$ für $l < N^{1/2}$ verteilt ist und favorisieren auf der Grundlage einer Simulationsstudie $l = N^4$.

A.5 Hinich Bispektraltest

Der Hinich Bispektraltest (1982) [62] untersucht genau wie der Tsay Test die dritten Momente einer Zeitreihenverteilung. Allerdings setzt der Test direkt bei den Zeitreihenwerten x_t und nicht bei den Fehlertermen u_t an und bildet zudem die dritten Momente mit Hilfe einer Fouriertransformation im Frequenzbereich ab. Das Bispektrum einer Zeitreihe ist somit die Fouriertransformation seiner dritten Momente. Bezeichnet man mit $c_{xxx}(r, s)$ die dritten Momente der Zeitreihe x_t :

$$c_{xxx}(r, s) = E[x_t x_{t-r} x_{t-s}] \quad (\text{A.10})$$

dann ist das Bispektrum für die Frequenzen f_1 und f_2 wie folgt definiert:

$$B_{xxx}(f_1, f_2) = \sum_{r=-\infty}^{\infty} \sum_{s=-\infty}^{\infty} c_{xxx}(r, s) \exp[-i2\pi(f_1 r + f_2 s)] \quad (\text{A.11})$$

und beschreibt eine doppelte Fouriertransformation der dritten Momente. Die Schiefefunktion $\Psi^2(f_1, f_2)$ wird wiederum mit Hilfe des Bispektrums definiert:

$$\Psi^2(f_1, f_2) = \frac{|B_{xxx}(f_1, f_2)|^2}{S_x(f_1) * S_x(f_2) * S_x(f_1 + f_2)} = \frac{\mu_3^2}{\sigma^6} \quad (\text{A.12})$$

mit S_x als Spektraldichtefunktion von x_t mit Frequenz f . Handelt es sich um einen linearen Prozeß, dann ist die Schiefefunktion $\Psi^2(f_1, f_2)$ für alle Frequenzbereiche konstant. Die Nullhypothese testet entsprechend auf einen konstanten Wert von $\Psi^2(f_1, f_2)$.

A.6 BDS Test

Ein weiterer Test, der aus dem Bereich der Chaos Theorie entspringt, wurde von Brock, Dechert und Scheinkman (1987) entwickelt. Diese nichtparametrische BDS Methode testet die aus dem Prewhitening Prozeß erhaltenen Fehlerterme u_t auf Unabhängigkeit. Dem Test liegt die folgende Idee zugrunde: Betrachtet man eine Zeitreihe mit identisch und unabhängig verteilten Beobachtungswerten, dann füllen diese Werte einen n-dimensionalen Raum gleichmäßig aus. Die Vorhersage eines Beobachtungswertes ist nicht möglich. Wird eine Zeitreihe hingegen durch einen nichtlinearen Datengenerierungsprozeß erzeugt, so projiziert der Prozeß ab einer bestimmten Raumdimension N eine geometrische Struktur (z.B.

Spirale oder Cluster). Für die Erzeugung von Strukturen müssen sich die Beobachtungswerte im Raum verdichten und treten somit im Vergleich zu einer identisch unabhängigen (Zufalls-)Verteilung häufiger auf. Die Nullhypothese der BDS Methode testet daher auf eine unabhängige und identische Verteilung und ist in der Lage, Nichtlinearitäten sowohl im Mittelwert als auch in der Varianz aufzudecken.

Der BDS Test basiert auf dem Konzept des Korrelationsintegrals nach Grassberger und Procaccia (1983):

$$C_N(\epsilon) = \sum_{1 \leq t < s \leq T_N}^N I(x_t^N, x_s^N) \left\{ \frac{2}{T_N(T_N - 1)} \right\} \quad (\text{A.13})$$

mit $T_N = T - N + 1$ und $I(x_t^N, x_s^N)$ als Indikatorfunktion. Betrachtet man eine Zeitreihe x_t mit T Beobachtungswerten, dann bezeichnet der n -dimensionale Vektor $x_t^N = (x_t, x_{t+1}, \dots, x_{t+N-1})$ die aufeinanderfolgenden Realisationen in einem n -dimensionalen Raum⁴. Diese sogenannte „N-History“ würde für identisch und unabhängige Zufallswerte einen n -dimensionalen Raum aufspannen, deren Werte im Raum gleichmäßig verteilt wären. Um zu überprüfen, ob die so dargestellten Werte Verdichtungen aufweisen, werden die Abstände zweier N-Histories x_t^N und x_s^N berechnet⁵. Die Indikatorfunktion nimmt den Wert eins (null) an, wenn der Abstand kleiner (größer) einem vorgegebenen Wert ϵ ist:

$$I(x_t^N, x_s^N) = \begin{cases} -1 & \text{wenn } \sup |x_{t+i} - x_{s+i}| \leq \epsilon \\ 0 & \text{sonst} \end{cases} \quad (\text{A.14})$$

Der BDS Test verwendet das Korrelationsintegral zusammen mit dem Dimensionsparameter m und dem Abstandsparameter ϵ und berechnet sich wie folgt:

$$w_{m,T}(\epsilon) = \sqrt{T - m + 1} * \frac{C_{m,T}(\epsilon) * C_{1,T-m+1}(\epsilon)^m}{\sigma_{m,T}(\epsilon)} \quad (\text{A.15})$$

Unter der Nullhypothese eines identisch und unabhängig verteilten Prozesses folgt der BDS Test asymptotisch einer Standardnormalverteilung:

$$\lim_{n \rightarrow \infty} w_{m,T}(\epsilon) \sim N(0, 1) \quad (\text{A.16})$$

Da der BDS Test für große Dimensionen große Datenmengen benötigt, um unverzerrte Teststatistiken zu erhalten, beschränkt sich die Anwendung auf die Dimensionen $m = 2, 3$ und 4.

⁴Beim graphisch maximal darstellbarem Raum mit drei Dimensionen geben die Spalten des transponierten Vektors x_t die jeweiligen Koordinaten an.

⁵Die Supremum Norm entspricht der Länge eines Vektors: $\sup|x| = \|x\| = \sqrt{x_1^2 + x_2^2 + \dots}$

Zeitreihe	Erhebungszeitraum	Signifikanzniveau ¹
Bundesrepublik Deutschland		
Null-Kupon Anleihe RLZ 5 Jahre	Januar 1973 - Dezember 2000	0.017
Null-Kupon Anleihe RLZ 4 Jahre	Januar 1973 - Dezember 2000	0.050
Null-Kupon Anleihe RLZ 3 Jahre	Januar 1973 - Dezember 2000	0.055
Null-Kupon Anleihe RLZ 2 Jahre	Januar 1973 - Dezember 2000	0.015
Null-Kupon Anleihe RLZ 1 Jahre	Januar 1973 - Dezember 2000	0.018
Geldmarktpapier	Januar 1973 - Dezember 2000	0.000
Inflationsrate*	Januar 1973 - Dezember 2000	0.001
Wechselkurs DM/US-\$	Januar 1974 - Dezember 1998	0.000
Geldmengenaggregat M1	Januar 1974 - Dezember 1998	0.000
Industrieproduktion	Januar 1974 - Dezember 1998	0.000
Zinssatz	Januar 1974 - Dezember 1998	0.003
Handelsbilanz	Januar 1974 - Dezember 1998	0.000
Niederlande		
Inflationsrate*	Januar 1974 - Dezember 1998	0.001
Wechselkurs Dutch Guilder/US-\$	Januar 1974 - Dezember 1998	0.000
Geldmengenaggregat M1	Januar 1974 - Dezember 1998	0.100
Industrieproduktion	Januar 1974 - Dezember 1998	0.000
Zinssatz*	Januar 1974 - Dezember 1998	0.000
Handelsbilanz	Januar 1974 - Dezember 1998	0.131
Japan		
Inflationsrate*	Januar 1974 - Dezember 2002	0.058
Wechselkurs Yen/US-\$	Januar 1974 - Dezember 2002	0.000
Geldmengenaggregat M1	Januar 1974 - Dezember 2002	0.130
Industrieproduktion	Januar 1974 - Dezember 2002	0.000
Zinssatz*	Januar 1974 - Dezember 2002	0.000
Handelsbilanz	Januar 1974 - Dezember 2002	0.969
Großbritannien		
Inflationsrate*	Januar 1973 - Dezember 2002	0.000
Wechselkurs £/US-\$	Januar 1974 - Dezember 2002	0.000
Geldmengenaggregat M0	Januar 1974 - Dezember 2002	0.279
Industrieproduktion	Januar 1974 - Dezember 2002	0.805
Zinssatz*	Januar 1974 - Dezember 2002	0.000
Handelsbilanz	Januar 1974 - Dezember 2002	0.000
U.S.A.		
Inflationsrate	Januar 1973 - Dezember 2002	0.000
Geldmengenaggregat M1	Januar 1974 - Dezember 2002	0.000
Industrieproduktion	Januar 1974 - Dezember 2002	0.000
Zinssatz	Januar 1974 - Dezember 2002	0.000
Handelsbilanz	Januar 1974 - Dezember 2002	0.000

Tabelle A.1: BDS Test nach Kanzler. ¹Test auf identisch unabhängig verteilte Residuen. Auswertung von Niveauvariablen bzw. Variablen in 1.Differenzen; Niveauvariablen sind mit einem * gekennzeichnet; m=2, eps=1, Anzahl Bootstrap-Simulationen = 25000.

Zeitreihe	Erhebungszeitraum	McLeod-Li	Engle LM	BDS	Tsay	Bicovariance	Bispectral
Bundesrepublik Deutschland							
Null-Kupon Anleihe RLZ 5 Jahre	Januar 1973 - Dezember 2000	0.001	0.001	0.001	0.001	0.002	0.033
Null-Kupon Anleihe RLZ 4 Jahre	Januar 1973 - Dezember 2000	0.001	0.002	0.000	0.000	0.002	0.012
Null-Kupon Anleihe RLZ 3 Jahre	Januar 1973 - Dezember 2000	0.001	0.001	0.000	0.000	0.002	0.007
Null-Kupon Anleihe RLZ 2 Jahre	Januar 1973 - Dezember 2000	0.002	0.001	0.000	0.002	0.001	0.050
Null-Kupon Anleihe RLZ 1 Jahre	Januar 1973 - Dezember 2000	0.000	0.000	0.000	0.006	0.000	0.094
Geldmarktpapier	Januar 1973 - Dezember 2000	0.000	0.000	0.000	0.011	0.002	0.210
Inflationsrate*	Januar 1973 - Dezember 2000	0.000	0.000	0.000	0.000	0.000	0.476
Wechselkurs DM/US-\$	Januar 1974 - Dezember 1998	0.000	0.004	0.239	0.014	0.000	0.005
Geldmengenaggregat M1	Januar 1974 - Dezember 1998	0.457	0.581	0.000	0.285	0.005	0.035
Industrieproduktion	Januar 1974 - Dezember 1998	0.016	0.001	0.001	0.003	0.010	0.439
Zinssatz	Januar 1974 - Dezember 1998	0.000	0.002	0.000	0.000	0.000	0.082
Handelsbilanz	Januar 1974 - Dezember 1998	0.000	0.022	0.034	0.000	0.000	0.757
Niederlande							
Inflationsrate*	Januar 1974 - Dezember 1998	0.000	0.175	0.550	0.546	0.695	0.580
Wechselkurs Dutch Guilder/US-\$	Januar 1974 - Dezember 1998	0.000	0.000	0.169	0.000	0.000	0.060
Geldmengenaggregat M1	Januar 1974 - Dezember 1998	0.000	0.008	0.023	0.000	0.001	0.051
Industrieproduktion	Januar 1974 - Dezember 1998	0.008	0.179	0.016	0.418	0.162	0.435
Handelsbilanz	Januar 1974 - Dezember 1998	0.000	0.013	0.015	0.821	0.000	0.183
Zinssatz*	Januar 1974 - Dezember 1998	0.549	0.117	0.000	0.013	0.090	0.176

Tabelle A.2: Die Tabelle zeigt die **p-Werte** der entsprechenden Teststatistiken; Auswertung von Niveauvariablen bzw. Variablen in 1. Differenzen; stationäre Niveauvariablen sind mit einem * gekennzeichnet; Nichtlineare Testverfahren

Zeitreihe	Erhebungszeitraum	McLeod-Li L=24	Engle LM p=5	BDS m=2	Tsay k=5	Bicovariance l=8	Bispectral M=24
Japan							
Inflationsrate*	Januar 1974 - Dezember 2002	0.000	0.018	0.000	0.000	0.000	0.893
Wechselkurs Yen/US-\$	Januar 1974 - Dezember 2002	0.045	0.086	0.003	0.046	0.002	0.038
Geldmengenaggregat M1	Januar 1974 - Dezember 2002	0.003	0.001	0.000	0.000	0.000	0.039
Industrieproduktion	Januar 1974 - Dezember 2002	0.018	0.044	0.000	0.166	0.003	0.753
Handelsbilanz	Januar 1974 - Dezember 2002	0.000	0.640	0.326	0.000	0.000	0.243
Zinssatz*	Januar 1974 - Dezember 2002	0.986	0.909	0.261	0.587	0.961	0.910
Großbritannien							
Inflationsrate*	Januar 1974 - Dezember 2002	0.002	0.036	0.093	0.000	0.025	0.069
Wechselkurs £/US-\$	Januar 1974 - Dezember 2002	0.000	0.000	0.000	0.000	0.000	0.554
Geldmengenaggregat M0	Januar 1974 - Dezember 2002	0.000	0.221	0.000	0.000	0.002	0.009
Industrieproduktion	Januar 1974 - Dezember 2002	0.001	0.002	0.000	0.017	0.000	0.475
Zinssatz*	Januar 1974 - Dezember 2002	0.000	0.000	0.000	0.013	0.001	0.881
Handelsbilanz	Januar 1974 - Dezember 2002	0.000	0.000	0.007	0.000	0.000	0.018
U.S.A.							
Inflationsrate*	Januar 1974 - Dezember 2002	0.020	0.006	0.000	0.090	0.011	0.820
Geldmengenaggregat M1	Januar 1974 - Dezember 2002	0.001	0.004	0.000	0.000	0.000	0.799
Industrieproduktion	Januar 1974 - Dezember 2002	0.385	0.059	0.004	0.141	0.890	0.645
Zinssatz	Januar 1974 - Dezember 2002	0.000	0.000	0.000	0.000	0.000	0.079
Handelsbilanz	Januar 1974 - Dezember 2002	0.000	0.000	0.210	0.003	0.000	0.000

Tabelle A.3: Die Tabelle zeigt die **p-Werte** der entsprechenden Teststatistiken; Auswertung von Niveauvariablen bzw. Variablen in 1.Differenzen; stationäre Niveauvariablen sind mit einem * gekennzeichnet; Nichtlineare Testverfahren

Anhang B

Symbolische Regression

B.1 Ergebnisse deterministisch dynamischer Prozesse

In diesem Anhang sind die Ergebnisse der symbolischen Regressionen von GP GAUSS aufgeführt. Mit Hilfe der Regressionen von nichtlinearen Prozessen wird gezeigt, wie die Güte der In-Sample Schätzungen von GP GAUSS tendenziell abnimmt, je komplexer der zu regressierende Prozeß ist.

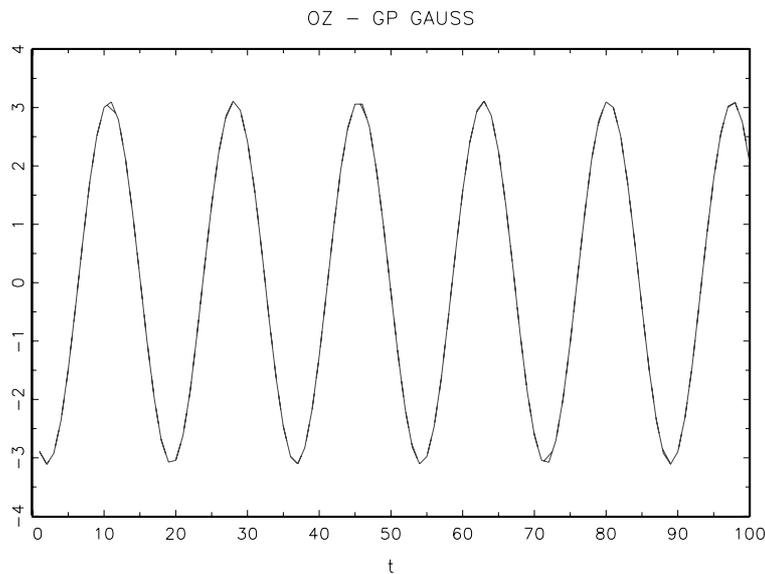


Abbildung B.1: Dynamischer Prozeß: OZ - beste GP GAUSS Regression

B.1.1 Lineare Ozaki Funktion

Als erstes wird GP GAUSS auf eine einfache lineare Funktion angewendet, die in einem Aufsatz von Ozaki [100] behandelt wird. In Abbildung B.1 ist sowohl die originale als auch

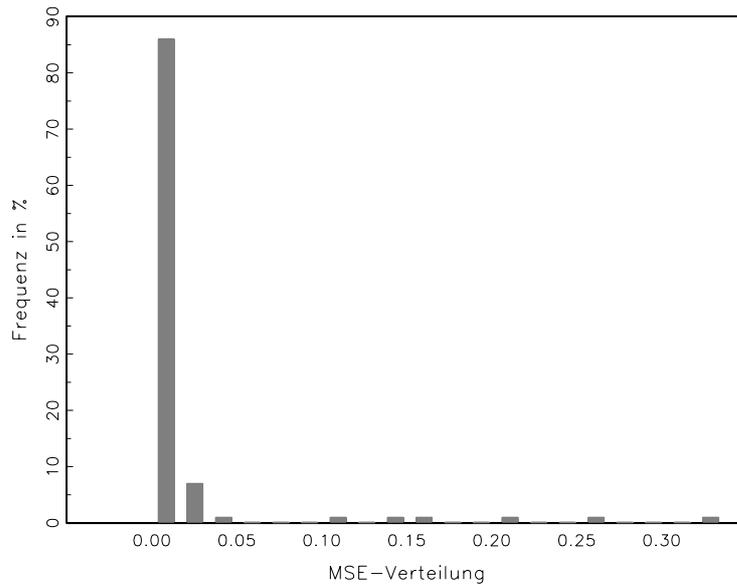


Abbildung B.2: Histogramm OZ - die besten MSE Werte aus 100 Läufen

die bestgeschätzte Zeitreihe wiedergegeben. Wegen einer Fehlerabweichung von beinahe null, wie man der Tabelle B.1 entnehmen kann, können die beiden Zeitreihen allerdings kaum unterschieden werden. Der Jarque-Bera Test auf Normalverteilung sowie der Box-Ljung Test auf Autokorrelation zeigen, daß die Fehlerabweichungen autokorreliert und nicht normalverteilt sind (Tabelle B.1).

Die beste von GP GAUSS entwickelte Regressionsgleichung sieht dabei folgendermaßen aus:

$$OZ = \frac{((Y_t[v - 0003]) - ((Y_t[v - 0006]) + ((Y_t[v - 0003]) / (wzl((wzl(((0007) + ((0058)))) * (ln(0007)))) * (ln(0049))))))}{(B.1)}$$

Die relativ kurze Schätzgleichung spiegelt den wenig komplexen Datengenerierungsprozeß wieder. Dies wird auch im Histogramm B.2 deutlich. Ungefähr 85% der genetisch entwickelten Regressionsgleichungen erweisen sich als sehr gute Regressionen mit einem MSE nahe null.

Test auf Normalverteilung		Box-Ljung-Test		Fitneßkriterien	
Kurtosis:	-1.520	Signifikanz - Lag 1:	0.000	SSE:	0.000002443
Schiefe:	-0.067	Signifikanz - Lag 5:	0.000	MSE:	0.000000024
Jarque-Bera:	9.704	Signifikanz - Lag 15:	0.000	NMSE:	0.000000005
Signifikanz JB:	0.008	Signifikanz - Lag 20:	0.000	R^2 :	0.999999995

Tabelle B.1: OZ Residuenanalyse

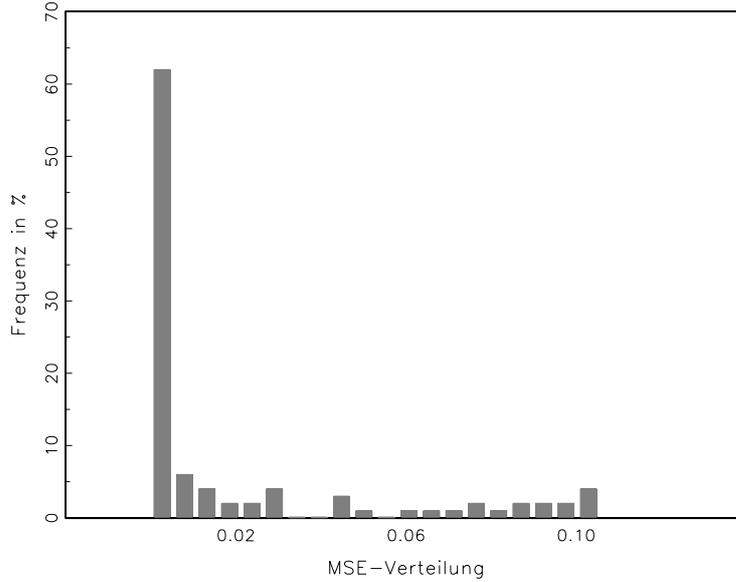


Abbildung B.4: Histogramm LG - die besten MSE Werte aus 100 Läufen

Test auf Normalverteilung		Box-Ljung-Test		Fitneßkriterien	
Kurtosis:	2.459	Signifikanz - Lag 1:	0.639	SSE:	0.000253319
Schiefe:	1.128	Signifikanz - Lag 5:	0.651	MSE:	0.000002533
Jarque-Bera:	46.407	Signifikanz - Lag 15:	0.959	NMSE:	0.000020512
Signifikanz JB:	0.000	Signifikanz - Lag 20:	0.989	R^2 :	0.999979206

Tabelle B.2: LG Residuenanalyse

B.1.3 Henon Map

Die Häufigkeitsverteilung der Henon Map (Henon [61]) spiegelt die zunehmende Komplexität der nichtlinearen chaotischen Funktion wieder. Der Anteil sehr guter Regressionsgleichungen reduziert auf ungefähr 36%, wenngleich die Regressionsgüte der besten Schätzgleichung mit einem MSE-Wert von 0.00066 weiterhin eine Fehlerabweichung von nahe null belegt. Die in Tabelle B.3 dargestellten Testergebnisse zeigen nicht normalverteilte und nicht-autokorrelierte Fehlerabweichungen. Die von GP GAUSS entwickelte Regressionsgleichung mit der geringsten Fehlerabweichung sieht folgendermaßen aus:

$$\begin{aligned}
 HN = & \left((wzl(abs(wzl(wzl(exp(((Y_t[v - 0002]) + (Y_t[v - 0002]) - (abs(\quad (B.3) \right. \\
 & (((Y_t[v - 0001]) * (wzl(cos(0.86)))) * ((wzl(((Y_t[v - 0001])) - \\
 & (abs((((0.04) - (cos(0.22))) - (abs((((((((((ln1((Y_t[v - 0001])) \\
 & +(cos((0.33)))) * ln1((((Y_t[v - 0001])) * ln1(wzl(((Y_t[v - 0001])) \\
 &)))) + (0.82)))))) + (0.84)))) * ln1((Y_t[v - 0001])) + (cos(0094)))) * \\
 & (ln1(0.04) + (cos(0000)))))))))) * ln1((0.42) + ((Y_t[v - 0002]) \\
 &)))))))) - (abs((((0.42) - (wzl((Y_t[v - 0001])))) - (abs((((
 \end{aligned}$$

$$(((Y_t[v - 0001])) * (\ln1(wzl((Y_t[v - 0001])) + (0.86)))) * (\ln1(Y_t[v - 0001] + (0.86))))))$$

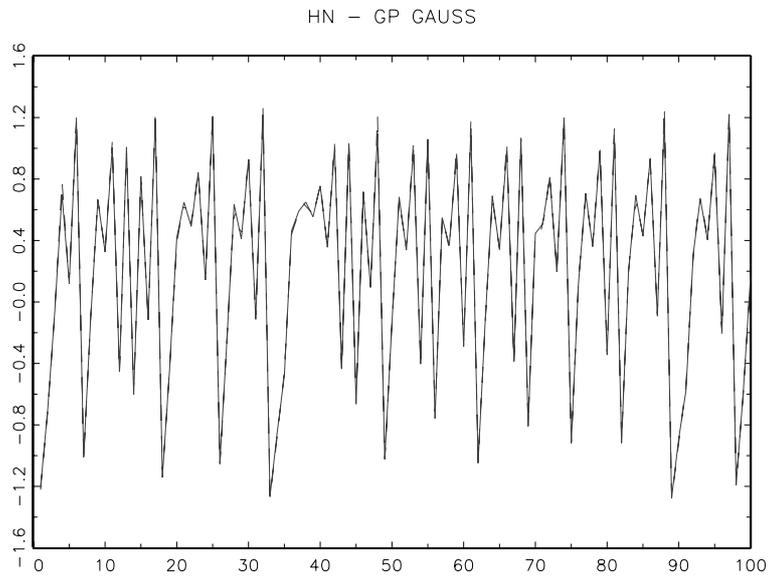


Abbildung B.5: Dynamischer Prozeß: HN - beste GP GAUSS Regression

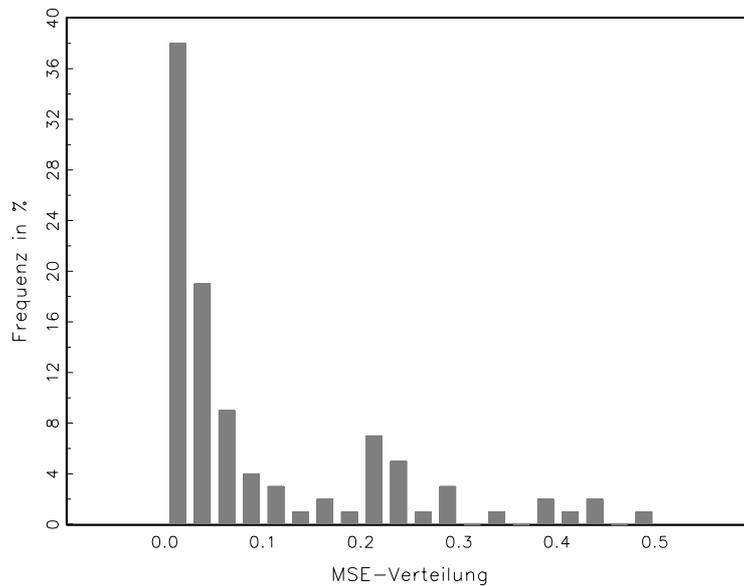


Abbildung B.6: Histogramm HN - die besten MSE Werte aus 100 Läufen

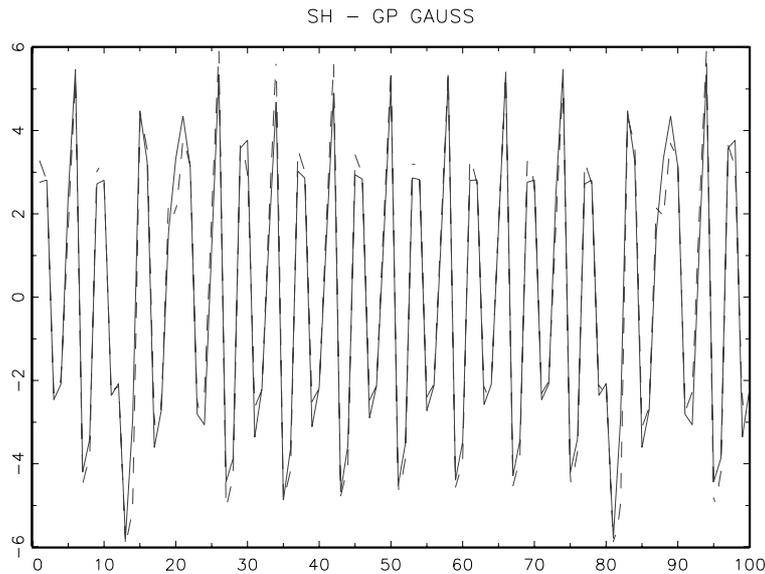


Abbildung B.7: Dynamischer Prozess: SH - beste GP GAUSS Regression

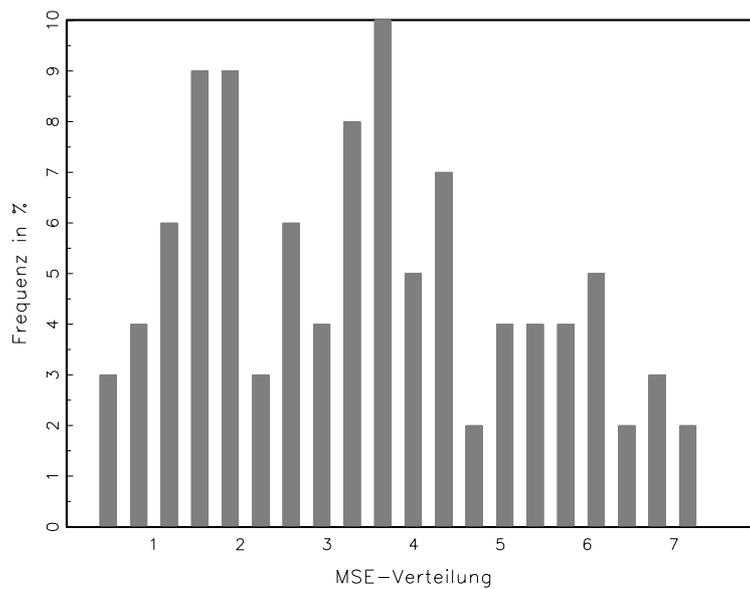


Abbildung B.8: Histogramm SH - die besten MSE Werte aus 100 Läufen

B.1.5 Mackey-Glass

Die Ergebnisse zur nichtlinearen chaotischen Funktion von Mackey-Glass [83] fallen etwas schlechter aus. Im Vergleich zum Genetischen Programm von Kaboudan TSGP besitzt die beste Regressionsgleichung von GP GAUSS eine etwas größere Fehlerabweichung. Der Jarque-Bera und Box-Ljung Test ergeben normalverteilte und nicht-autokorrelierte Fehlerabweichungen. Die beste von GP GAUSS entwickelte Regressionsgleichung sieht folgen-

dermaßen aus:

$$\begin{aligned}
 MG = & \left((wzl(wzl((((wzl((((wzl(((wzl((wzl((0063))) * \right. & \text{(B.5)} \\
 & (sin((wzl(sin((0.16) + (v) + (Y_t[v - 0005]))) * \\
 & (sin(wzl((0097)/(Y_t[v - 0007]) - (Y_t[v - 0002]) - \\
 & (ln1((Y_t[v - 0003]) * (wzl(Y_t[v - 0004]) * (0.03)/ \\
 & (wzl(Y_t[v - 0011]) + (ln1((((0092)))))))))) \\
 &)))))))) * ((wzl(sin((Y_t[v - 0002]) * (wzl(\\
 & ((Y_t[v - 0010]) * (wzl((((0038)))))))))) \\
 &)))) * (wzl(((wzl(Y_t[v - 0001])) * (wzl(((wzl \\
 & (wzl(abs(((Y_t[v - 0009]) * (Y_t[v - 0003])) * (((wzl(\\
 & 0.11) * (sin((Y_t[v - 0002]) * (wzl((0066)))))))))) \\
 &)) * (((wzl(sin(((Y_t[v - 0004]) + (exp(((wzl((0.19) \\
 &)))) * (wzl(((Y_t[v - 0005])/(0.11)))))) * (sin((\\
 & Y_t[v - 0002]) + ((Y_t[v - 0003]) - (wzl((0006) + (wzl((\\
 & ((Y_t[v - 0011])) * (ln1((Y_t[v - 0006])))))))))))) * \\
 & (((((((wzl(((wzl(0097))))))))))))))
 \end{aligned}$$

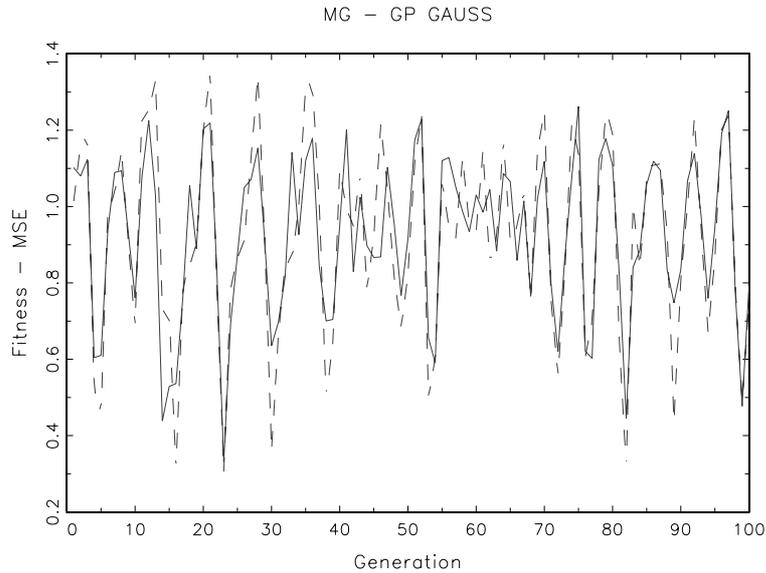


Abbildung B.9: Dynamischer Prozeß: MG - beste GP GAUSS Regression

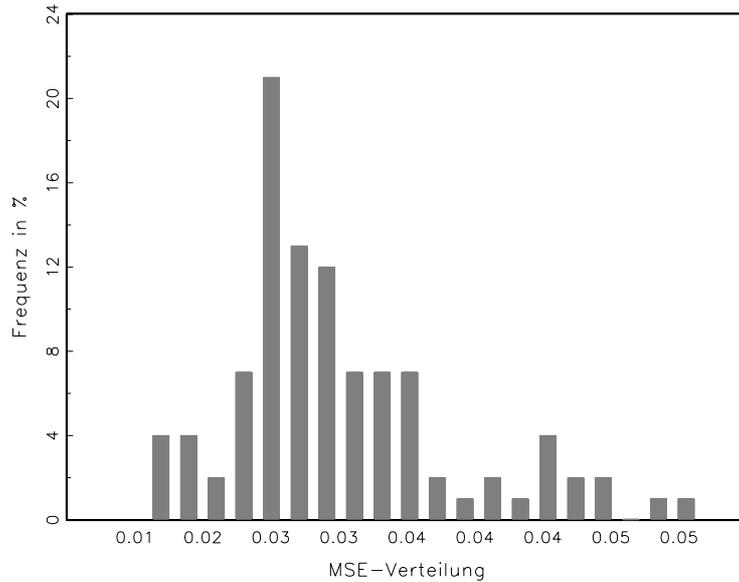


Abbildung B.10: Histogramm MG - die besten MSE Werte aus 100 Läufen

Test auf Normalverteilung		Box-Ljung-Test		Fitneßkriterien	
Kurtosis:	0.062	Signifikanz - Lag 1:	0.870	SSE:	1.590254511
Schiefe:	-0.068	Signifikanz - Lag 5:	0.441	MSE:	0.015902545
Jarque-Bera:	0.093	Signifikanz - Lag 15:	0.778	NMSE:	0.241648030
Signifikanz JB:	0.954	Signifikanz - Lag 20:	0.540	R^2 :	0.735063963

Tabelle B.5: MG Residuenanalyse

B.2 Ergebnisse deterministisch stochastischer Prozesse

Bei den symbolischen Regressionen zuvor handelte es sich um nichtlineare chaotische Datengenerierungsprozesse ohne stochastische Elemente. Im folgenden werden zwei nichtlineare stochastische Prozesse betrachtet.

B.2.1 Bilineares System von Granger & Andersen

Vergleicht man die bisherigen Häufigkeitsverteilungen einschließlich des Bilinearen Systems von Granger & Andersen [53] miteinander, so erkennt man eine zunehmende Streuung in der MSE Verteilung. Die in Tabelle B.3 dargestellten Testergebnisse zeigen nicht normalverteilte und nicht-autokorrelierte Fehlerabweichungen. Die von GP GAUSS entwickelte Regressionsgleichung mit der geringsten Fehlerabweichung sieht folgendermaßen aus:

$$\begin{aligned}
 BL = & \left((wzl(0.35) * (sin(abs(0.25) * (Y_t[v - 0001]))) * ((((((Y_t[v - 0001])))) - (0.95)) - (cos(((v) * (cos((0052)))) - (((Y_t[v - 0011]) + ((0026)))))))))) + \right. \\
 & \left. (sin(abs(sin(.23) * (((Y_t[v - 0001])) - (exp((sin((v) - (Y_t[v - 0009])) * (
 \end{aligned}
 \tag{B.6}$$

$$\begin{aligned}
& Y_t[v - 0001]) - ((v) + (\ln(Y_t[v - 0008]) + (Y_t[v - 0008]))) + (((0041)))))) \\
&) * ((0.45) + (Y_t[v - 0005])) + (v) + (((0002) + (\text{abs}((((((((0.35)))))))))))/ \\
& Y_t[v - 0009]) - (\ln((Y_t[v - 0011]) + (0023)))))) * (\sin(\text{abs}((0038) + (\cos \\
& ((Y_t[v - 0010]) - (\cos((0.61))))))) * (Y_t[v - 0001])))) + (\sin((v) * (\ln(0.33))) \\
& *(Y_t[v - 0001]))
\end{aligned}$$

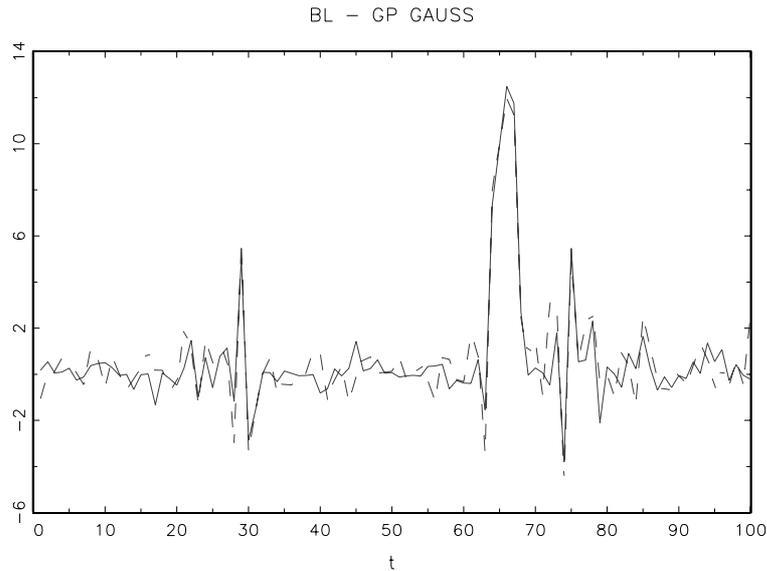


Abbildung B.11: Dynamischer Prozeß: BL - beste GP GAUSS Regression

Test auf Normalverteilung		Box-Ljung-Test		Fitneßkriterien	
Kurtosis:	1.318	Signifikanz - Lag 1:	0.815	SSE:	88.782245963
Schiefe:	-0.613	Signifikanz - Lag 5:	0.735	MSE:	0.887822460
Jarque-Bera:	13.491	Signifikanz - Lag 15:	0.688	NMSE:	0.145482489
Signifikanz JB:	0.001	Signifikanz - Lag 20:	0.476	R^2 :	0.860477009

Tabelle B.6: BL Residuenanalyse

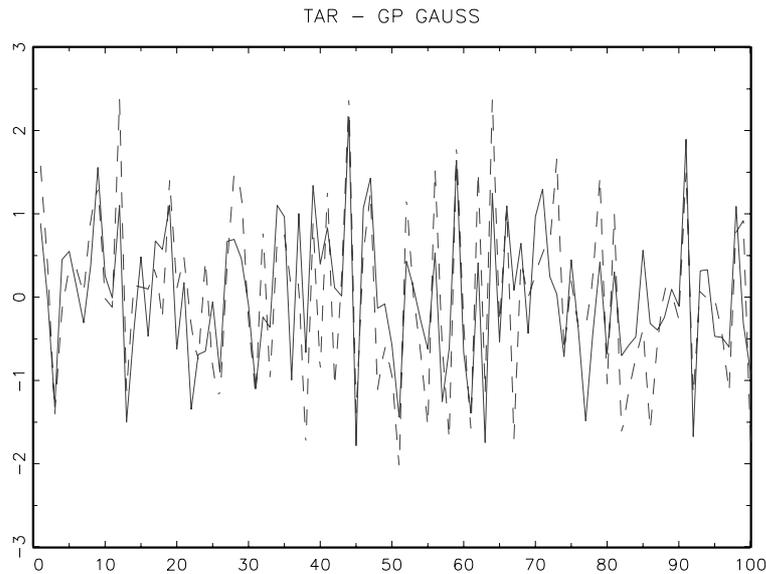


Abbildung B.13: Dynamischer Prozeß: TAR - beste GP GAUSS Regression

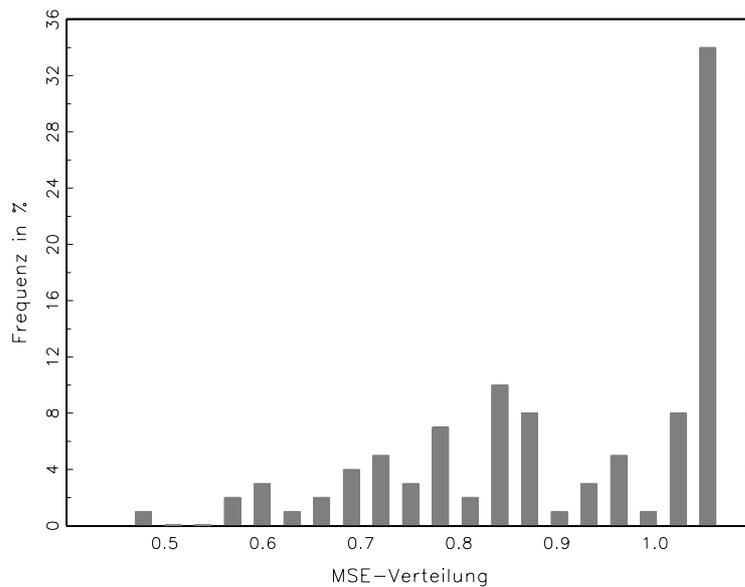


Abbildung B.14: Histogramm TAR - die besten MSE Werte aus 100 Läufen

B.3 Ergebnisse bei Schätzen von Pseudo-Zufallszahlen

In der letzten symbolischen Schätzung sollen Zufallszahlen regressiert werden. Wie man der Tabelle B.8 entnehmen kann, liefert die beste Regression erstaunlich gute Werte mit einer geringen Fehlerabweichung. Ein Blick auf das zugehörige Histogramm zeigt, daß diese Regression mit einem MSE Wert von 11.7 aus insgesamt einhundert Läufen eine Ausnahme darstellt. Die zweitbeste Regression besitzt bereits eine wesentlich größere Fehlerab-

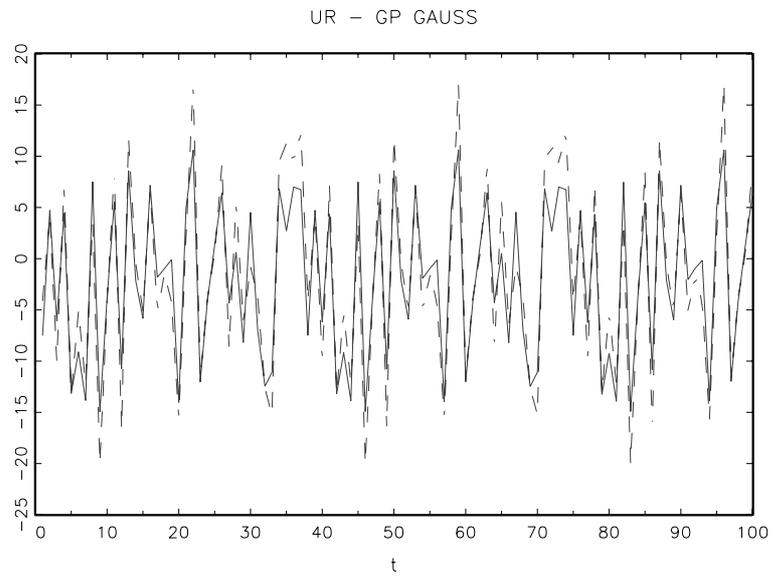


Abbildung B.15: Dynamischer Prozess: UR - beste GP GAUSS Regression

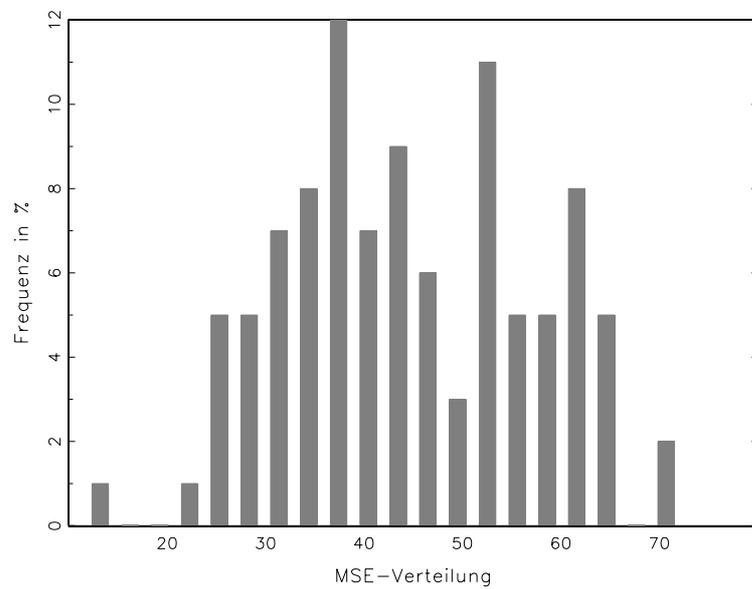


Abbildung B.16: Histogramm UR - die besten MSE Werte aus 100 Läufen

Anhang C

Weitere Auswertungen der Inflationsprognosemodelle

C.1 Schätzung von Zinsstrukturkurven - Svensson Methode

Die folgenden Erläuterungen zur Schätzung von Zinsstrukturkurven orientieren sich an dem Monatsbericht der Bundesbank von Oktober 1997. Die Zinsstrukturkurve zeigt den graphischen Zusammenhang zwischen Zinssätzen und Laufzeiten ohne Kreditausfallrisiko. Bei der Betrachtung von Null-Kupon-Anleihen ist die Zinsberechnung aufgrund der einmaligen Zahlung sehr einfach. Die (jährliche) Ertragsrate berechnet sich aus Rückzahlungswert und aktuellem Kurs. Da die überwiegende Zahl der Anleiheemissionen während ihrer Laufzeit mehrere Kuponzahlungen beinhalten, müssen diese Zahlungsströme mit dem Zinssatz abdiskontiert werden, der nach den gegenwärtigen Marktverhältnissen zu erwarten ist. Für die Schätzung von Zinsstrukturkurven verwendet die Deutsche Bundesbank einen von Nelson und Siegel [96] konzipierten und von Svensson [118] erweiterten Ansatz, der folgenden funktionalen Zusammenhang zwischen Zinssätzen und Laufzeiten annimmt. Demnach wird der Zinssatz als Summe aus einer Konstanten und verschiedenen Exponentialtermen sowie als Funktion von insgesamt sechs Parametern definiert:

$$z(T, \beta) = \beta_0 + \beta_1 * \left(\frac{1 - \exp(-T/\tau_1)}{T/\tau_1} \right) \quad (\text{C.1})$$

$$+ \beta_2 * \left(\frac{1 - \exp(-T/\tau_2)}{T/\tau_1} - \exp(-(T/\tau_1)) \right) \quad (\text{C.2})$$

$$+ \beta_3 * \left(\frac{1 - \exp(-T/\tau_2)}{T/\tau_2} - \exp(-T/\tau_2) \right) \quad (\text{C.3})$$

Der Term $z(T, \beta)$ beschreibt den Zinssatz für Laufzeit T als Funktion des Parametervektors β und die Variablen $\beta_0, \beta_1, \beta_2, \beta_3, \tau_1$ und τ_2 die zu schätzenden Parameter dieses Vektors. Die Erweiterung des Ansatzes durch Svensson besteht in der Berücksichtigung des letzten Terms, der einen zusätzlichen Wendepunkt der geschätzten Kurve erlaubt. Die langfristig extrapolierten Zinssätze konvergieren gegen den Wert β_0 , da die Exponentialterme mit zunehmender Laufzeit gegen null gehen. Für die Schätzungen werden die von Bundesanleihen, -obligationen und -schatzanweisungen mit (Rest-)Laufzeiten von mindestens

drei Monaten zugrundegelegt. Die Parameter werden anhand eines nichtlinearen Optimierungsverfahrens geschätzt. Die Optimierung erfolgt durch die Minimierung der quadrierten Abweichungen zwischen geschätzten und beobachteten Renditen. Die Aufstellung von Nebenbedingungen für einige Parameter gewährleistet, daß die geschätzten Zinssätze positiv sind.

C.2 ARIMA

Zur Bestimmung der Modellspezifikation wird das Akaike und das Schwarz Informationskriterium für verschiedene autorregressive und Moving Average Grade p und q berechnet und die Modellordnung mit dem kleinsten Wert gewählt. Danach favorisiert das Akaike

		q			
		0	1	2	3
p	0		506.00	507.96	501.70
	1	520.57	507.98	506.38	502.38
	2	495.97	496.17	495.20	494.22
	3	495.40	497.26	495.85	495.57

Tabelle C.1: AIC für deutsche Inflationsraten 1973-1996

		q			
		0	1	2	3
p	0		513.32	518.94	516.33
	1	527.89	518.96	521.02	520.68
	2	506.95	510.81	513.50	516.18
	3	510.03	515.56	517.81	521.19

Tabelle C.2: SIC für deutsche Inflationsraten 1973-1996

Informationskriterium ein ARIMA(2,1,3) und das Schwarz Informationskriterium ein ARIMA(2,1,0). Der Differenzegrad für die Inflationsraten $d = 1$ wurde bereits bei der Analyse der Zeitreiheneigenschaften anhand des erweiterten Dickey-Fuller Tests (ADF-Test) berechnet (Tabelle 4.1). Um zu überprüfen, ob die Modelle korrekt spezifiziert sind, müssen die Schätzresiduen e auf Autokorrelation und Normalverteilung getestet werden. Die Tests erfolgen mit Hilfe des Box-Ljung (LBQ) und des Jarque-Bera (JB) Tests. Aus den beiden Tabellen C.3 und C.4 ist ersichtlich, daß weder die Schätzresiduen des ARIMA(2,1,0) noch die eines weniger streng parametrisierten ARIMA(2,1,3) Modells die Eigenschaften eines Weißen Rauschens aufweisen. Zwar sind die normalen Residuen in beiden Modellen unkorreliert, aber dafür besitzen ihre quadrierten Werte signifikante Autokorrelationswerte. Darüber hinaus lehnt der Jarque-Bera Test in beiden Fällen die Nullhypothese normalverteilter Prognosefehler eindeutig ab.

k	AC(e)	LBQ	Prob	AC(e ²)	LBQ	Prob		
1	0.03	0.22	0.64	0.16	7.76	0.01	Schiefe:	0.03
2	0.04	0.66	0.72	0.12	11.89	0.00	Kurtosis*:	6.45
3	0.06	1.73	0.63	0.57	106.52	0.00	JB:	497.21
4	-0.03	2.02	0.73	0.15	112.80	0.00	Prob:	0.00
5	-0.09	4.28	0.51	0.04	113.27	0.00		

Tabelle C.3: Analyse der Schätzresiduen für ARIMA(2,1,0); * Überschlußkurtosis

k	AC(e)	LBQ	Prob	AC(e ²)	LBQ	Prob		
1	0.01	0.02	0.89	0.13	4.54	0.03	Schiefe:	0.07
2	-0.04	0.44	0.80	0.09	7.06	0.03	Kurtosis*:	7.21
3	0.02	0.56	0.90	0.57	103.27	0.00	JB:	621.74
4	-0.04	1.10	0.89	0.13	108.01	0.00	Prob:	0.00
5	-0.02	1.19	0.95	0.02	108.08	0.00		

Tabelle C.4: Analyse der Schätzresiduen für ARIMA(2,1,3); * Überschlußkurtosis

C.3 ARFIMA

parameter	estimates	std.err.	t-statistic	p-value
ϕ_1	-0.022539	0.083186	-0.270943	0.786659
θ_1	0.762712	0.056466	13.507376	0.000000
d	0.379900	0.000000	.	.
σ	0.604307	0.026707	22.627133	0.000000

Tabelle C.5: Parameterschätzungen eines ARFIMA(1, 0.3799, 1) Modells für differenzierte Inflationsraten mit $n = 264$ und einem GPH Schätzwert von $d(0.5) = 0.3799$. Die optimale Anzahl der Lags für die AR und MA Parameter wurde bei einer maximal möglichen Modellordnung von $p = 1$ und $q = 1$ anhand des BIC Kriteriums berechnet.

parameter	estimates	std.err.	t-statistic	p-value
ϕ_1	0.020288	0.190125	0.106706	0.915105
θ_1	0.413570	0.149328	2.769539	0.006020
d	-0.029717	0.000000	.	.
σ	0.603730	0.018867	31.998668	0.000000

Tabelle C.6: Parameterschätzungen eines ARFIMA(1, d , 1) Modells für differenzierte Inflationsraten mit $n = 264$. Die optimale Anzahl der Lags für die AR und MA Parameter wurde bei einer maximal möglichen Modellordnung von $p = 1$ und $q = 1$ anhand des BIC Kriteriums berechnet.

C.4 Genetische Programme: weitere Ergebnisse zu Inflationsprognosen

Die nachfolgende Gleichung gibt die ex-post beste Prognosegleichung aus dem Abschnitt 4.4.1 wieder, wie sie dem Anwender von GP GAUSS original angezeigt wird:

$$\pi_t^{24} = (((\hat{\pi}^{24}[t - 0001])) - (\sin((\sin((v)) + (\sin((0040)) + (\sin(\pi^0[t - 0006])) + (\sin((v)) + (\sin((0040)) + (\sin((\pi^0[t - 0006])) + (\sin((i^2[t - 0009])) + (((0.69) - (\text{abs}((i^0[t - 0004]))/(0.75)))))))/(\hat{\pi}^{24}[t - 0002]) + (((((\pi^0[t - 0010]) - (\exp((((\pi^0[t - 0010])/(i^0[t - 0004])))))))) - (\hat{\pi}^{24}[t - 0001]) + (\pi^0[t - 0010])))))))))/((0040) - (\text{abs}((i^0[t - 0004]))/(wzl((\text{abs}((((\pi^0[t - 0010]) - (\text{abs}(((i^0[t - 0004]))/(0097)/((\pi^0[t - 0010])/(i^0[t - 0004])))) * ((i^0[t - 0004])) * (0090)))))))))) - (i^0[t - 0004])))))))) \quad (\text{C.4})$$

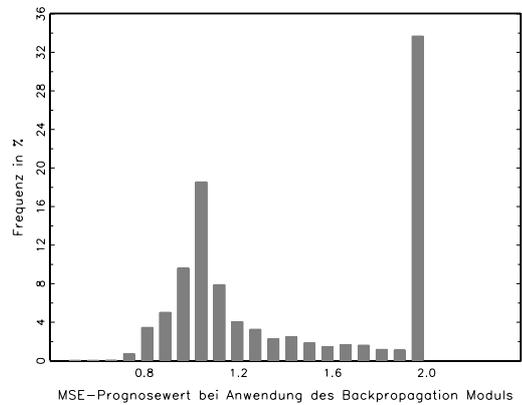
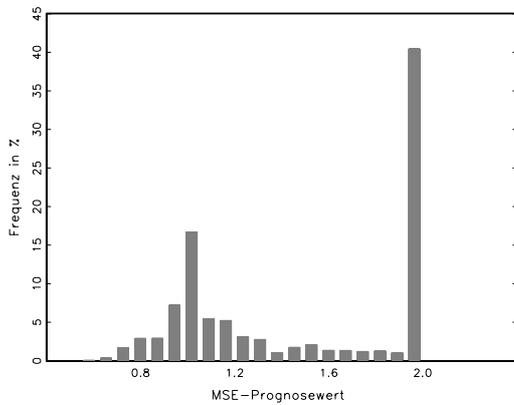
Die zahlreichen Klammersymbole sind für die genetischen Operationen notwendig, um insbesondere einen korrekten Ablauf bei den Crossovervorgängen sicherzustellen.

Modell	In-Sample	MSE Prognose	Sign.	Trefferquote	Sign.
beste Funktion	0.605	1.091	0.53	49.29	0.95
beste Funktion	0.659	1.413	—	49.30	0.96
besten 10 Funktionen	0.659	0.931	0.59	52.11	0.79
globales Minimum	0.705	0.908	0.59	53.52	0.69
globale 10 Minima	0.735	0.817	0.61	49.30	0.94

Tabelle C.7: 2.Simulationslauf auf der Basis eines **MSE Fitneßkriteriums**: verschiedene Auswahlstrategien. Pop = 100, Gen = 200, Runs = 100

Modell	In-Sample	MSE Prognose	Sign.	Trefferquote	Sign.
beste Funktion	-0.444	0.999	0.56	49.30	0.93
beste Funktion	-0.305	1.051	0.54	52.11	0.76
besten 10 Funktionen	-0.091	0.933	0.57	52.11	0.80
globales Minimum	-0.128	1.289	—	49.29	0.92
globale 10 Minima	-0.119	2.135	—	45.07	0.56

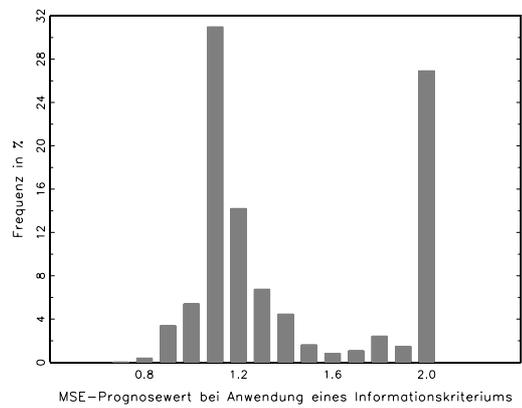
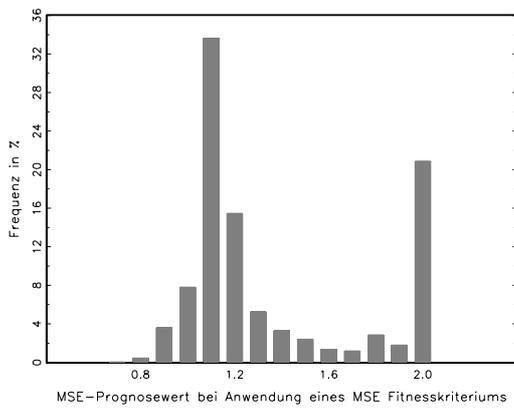
Tabelle C.8: 2.Simulationslauf auf der Basis eines **AIC Fitneßkriteriums**: verschiedene Auswahlstrategien. Pop = 100, Gen = 200, Runs = 100



(a) Inflationsprognose: In-Sample-, Trainings- und Out-of-Sample Menge

(b) Inflationsprognose: Backpropagation Modul

Abbildung C.1: Histogramme Inflationsprognosen: Abschnitt 4.4.2 und 4.4.3



(a) 2.Simulationslauf: Inflationsprognose auf der Basis eines MSE Fitneßkriteriums; Pop = 100, Gen = 200, Runs = 100

(b) 2.Simulationslauf: Inflationsprognose auf der Basis eines AIC Fitneßkriteriums; Pop = 100, Gen = 200, Runs = 100

Abbildung C.2: Histogramme Inflationsprognosen: Wiederholte GP GAUSS Simulation

Nr.	Lauf	In-Sample	Training	Out-of-Sample	InSmpl-RQ	InSmpl-RT	Trai-RQ	Trai-RT	Out-of-Smpl-RQ	Out-of-Smpl-RT
1	28	0.573	10.617	0.932	0.511	0.034	0.011	0.478	0.012	0.507
2	99	0.598	0.263	1.262	0.520	0.164	0.348	1.947	0.465	0.201
3	38	0.606	0.159	1.353	0.515	0.106	0.391	1.467	0.521	0.091
4	16	0.606	0.140	1.039	0.507	0.011	0.348	1.947	0.521	0.073
5	17	0.611	0.129	1.216	0.507	0.018	0.391	1.467	0.535	0.214
6	36	0.615	0.135	1.060	0.515	0.087	0.348	1.947	0.521	0.065
7	74	0.636	1.400	1.436	0.520	0.190	0.304	2.490	0.521	0.091
8	91	0.652	0.138	1.039	0.524	0.236	0.348	1.947	0.507	0.009
9	9	0.654	0.137	2.609	0.520	0.190	0.348	1.947	0.507	0.012
10	14	0.658	0.105	0.854	0.515	0.087	0.391	1.467	0.507	0.006
11	50	0.658	0.138	1.025	0.524	0.236	0.348	1.947	0.521	0.091
12	87	0.662	0.178	3.875	0.537	0.573	0.304	1.791	0.493	0.009
13	73	0.671	0.163	1.062	0.533	0.418	0.348	1.947	0.521	0.091
14	76	0.674	0.133	1.031	0.520	0.217	0.348	1.947	0.521	0.073
15	78	0.675	0.138	1.333	0.524	0.236	0.348	1.947	0.563	0.588

Tabelle C.9: Die Ergebnisse dieser Tabelle und der folgenden drei Tabellen beziehen sich auf einen weiteren Simulationslauf. Sie sollen dem Betrachter einen etwas detaillierteren Einblick vermitteln. In dieser Tabelle werden die Regressionsgleichungen anhand der besten In-Sample Schätzung ausgewählt. RQ = Trefferquote, RT = Teststatistik der Trefferquote.

Nr.	Lauf	Generation	Training	In-Smpl	Out-of-Smpl	InSmpl RQ	InSmpl RT	Training RQ	Training RT	Out-of-Smpl RQ	Out-of-Smpl RT
1	4	178	0.099	0.712	0.961	0.515	0.070	0.391	1.467	0.507	0.012
2	4	175	0.100	0.713	0.900	0.511	0.046	0.348	1.947	0.493	0.004
3	4	110	0.102	0.719	0.911	0.507	0.011	0.391	1.467	0.493	0.006
4	4	192	0.102	0.708	1.053	0.515	0.070	0.391	1.467	0.507	0.012
5	4	114	0.103	0.719	0.901	0.507	0.011	0.391	1.467	0.507	0.006
6	4	118	0.103	0.718	0.907	0.507	0.011	0.391	1.467	0.493	0.004
7	4	177	0.103	0.713	0.869	0.511	0.046	0.348	1.947	0.493	0.004
8	4	121	0.103	0.718	0.904	0.507	0.011	0.391	1.467	0.493	0.004
9	4	123	0.104	0.718	0.920	0.507	0.011	0.391	1.467	0.507	0.012
10	4	176	0.104	0.713	0.871	0.511	0.046	0.348	1.947	0.493	0.004
11	4	189	0.104	0.709	1.105	0.511	0.034	0.391	1.467	0.521	0.073
12	4	187	0.104	0.710	1.028	0.515	0.070	0.391	1.467	0.507	0.012
13	4	151	0.105	0.715	0.932	0.511	0.034	0.391	1.467	0.493	0.004
14	14	199	0.105	0.658	0.854	0.515	0.087	0.391	1.467	0.507	0.006
15	4	96	0.106	0.721	0.882	0.507	0.011	0.391	1.467	0.507	0.006
Out-of-Sample Durchschnitt für Nr. 1-10:					MSE = 0.920						

Tabelle C.10: Das Auswahlkriterium stellt in dieser Tabelle die besten MSE Prognosewerte aus der Trainingsperiode dar. Die Ergebnisse spiegeln ein typisches Ergebnis wieder. Nahezu sämtliche minimalen Prognosewerte aus der Trainingsperiode stammen aus dem gleichen Lauf (hier: Run Nr. 4).

Nr.	Lauf	Generation	Training	Fitneß	Prognose	Crossover	RichtInSm	Sign	RichTrai	Sign	RichtProg	Sign
1	87.000	11.000	0.118	0.764	1.167	0.901	0.520	0.708	0.391	0.226	0.465	0.173
2	58.000	24.000	0.118	0.777	1.633	0.648	0.511	0.782	0.304	0.115	0.521	0.073
3	43.000	12.000	0.123	0.781	1.066	0.777	0.520	0.708	0.348	0.163	0.507	0.006
4	98.000	30.000	0.124	0.780	1.074	0.403	0.520	0.708	0.348	0.163	0.507	0.006
5	13.000	17.000	0.124	0.765	1.315	0.858	0.520	0.708	0.348	0.163	0.521	0.091
6	49.000	23.000	0.125	0.777	2.024	0.957	0.515	0.768	0.348	0.163	0.521	0.102
7	26.000	13.000	0.125	0.758	1.077	0.414	0.520	0.708	0.348	0.163	0.507	0.006
8	18.000	14.000	0.125	0.780	1.078	0.862	0.520	0.708	0.348	0.163	0.507	0.006
9	42.000	8.000	0.125	0.784	1.096	0.161	0.511	0.854	0.348	0.163	0.507	0.012
10	69.000	22.000	0.126	0.759	1.079	0.747	0.520	0.708	0.348	0.163	0.507	0.006
11	74.000	15.000	0.127	0.778	1.003	0.968	0.520	0.708	0.348	0.163	0.507	0.006
12	6.000	49.000	0.129	0.718	0.995	0.652	0.520	0.685	0.348	0.163	0.507	0.006
13	17.000	17.000	0.129	0.705	1.248	0.958	0.520	0.685	0.391	0.226	0.521	0.081
14	22.000	30.000	0.129	0.786	1.098	0.976	0.520	0.708	0.348	0.163	0.521	0.091
15	20.000	10.000	0.129	0.781	1.068	0.702	0.524	0.627	0.348	0.163	0.507	0.006
Out-of-Sample Durchschnitt für Nr. 1-15:					MSE =	1.211						

Tabelle C.11: Die Auswahl der genetisch entwickelten Regressionsgleichungen erfolgt in dieser Tabelle wie folgt: Für jeden einzelnen Lauf wurden die aggregierten absoluten Fitnessänderungen der Crossover-Operationen in jeder Generation berechnet. Anschließend wurde für jeden Lauf diejenige Generation ermittelt, bei der die Fitnessänderungen die 1%-Grenze unterschritten hat. Aus dieser Generation wiederum wurde die Regressionsgleichung mit der minimalen In-Sample Fehlerabweichung gewählt. Die selektierten 100 Gleichungen wurden schließlich nach ihren minimalen Prognosefehlern (**MSE Punktprognosen**) aus der Trainingsperiode sortiert. Auswahlkriterium: Building Block < 1% - Trainingspunktprognosen.

Lauf	Generation	Knoten	Fitneß	Training	Prognose	RichtInSm	Sign	RichTrai	Sign	RichtProg	Sign
2	15	10	1.013	0.247	1.034	49.780		73.913	2.578	52.113	0.073
64	8	15	1.015	0.241	1.030	50.661	0.039	73.913	2.578	52.113	0.073
66	7	18	1.006	0.195	0.973	48.458	0.086	73.913	2.578	50.704	0.012
35	22	12	1.010	0.398	1.066	49.780	0.001	73.913	2.578	52.113	0.006
72	21	19	0.989	0.317	1.054	51.101	0.046	73.913	2.578	52.113	0.073
80	35	11	0.979	0.273	0.993	51.101	0.060	73.913	2.578	52.113	0.073
81	53	9	1.026	0.222	1.157	51.101	0.060	73.913	2.578	50.704	0.073
89	14	30	0.975	0.184	0.734	52.863	0.471	73.913	2.578	56.338	0.073
92	17	33	1.018	0.228	1.250	49.780	0.000	73.913	2.578	52.113	0.567
48	18	25	1.004	0.210	0.995	50.220	0.007	73.913	2.578	52.113	0.889
Out-of-Sample Durchschnitt für Nr. 1-10:					MSE =	1.155					

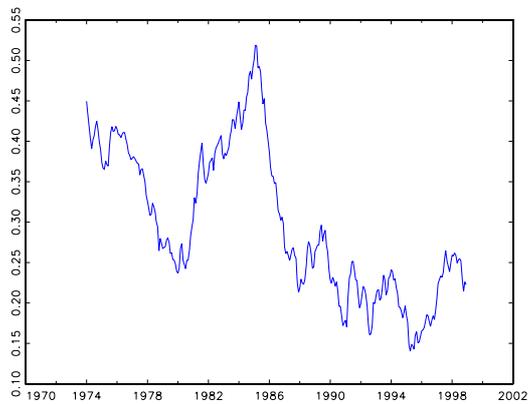
Tabelle C.12: Die Auswahl der genetisch entwickelten Regressionsgleichungen erfolgt in dieser Tabelle wie folgt: Für jeden einzelnen Lauf wurden die aggregierten absoluten Fitnessänderungen der Crossover-Operationen in jeder Generation berechnet. Anschließend wurde für jeden Lauf diejenige Generation ermittelt, bei der die Fitnessänderungen die 1%-Grenze unterschritten hat. Aus dieser Generation wiederum wurde die Regressionsgleichung mit der minimalen In-Sample Fehlerabweichung gewählt. Die selektierten 100 Gleichungen wurden schließlich nach ihren minimalen Prognosefehlern (**Richtungsprognosen**) aus der Trainingsperiode sortiert. Auswahlkriterium: Building Block < 1%.

Anhang D

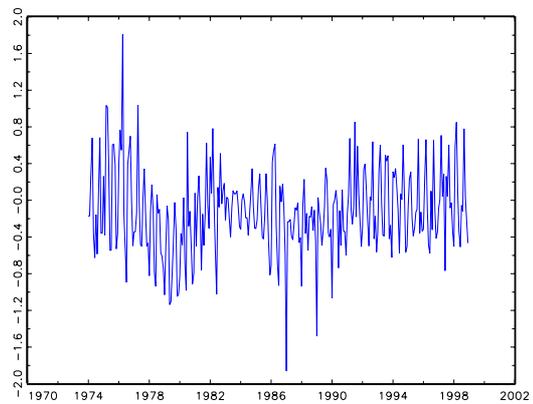
Weitere Auswertungen der Wechselkursmodelle

D.1 Daten für Wechselkursmodelle

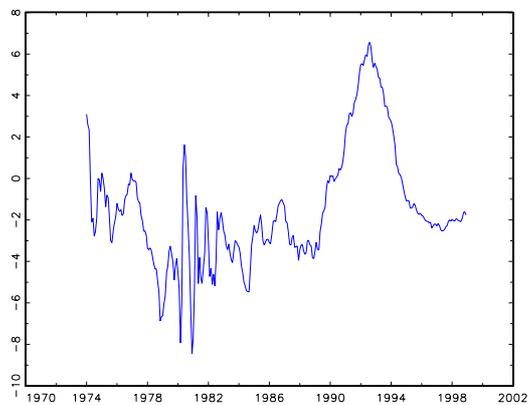
Für die empirischen Untersuchungen der Wechselkursmodelle wurden OECD Daten verwendet. Tabelle D.1 gibt eine exakte Beschreibung der einzelnen Zeitreihen wieder, wobei die Kodierungen der 5.Spalte für die DATASTREAM Datenbank gelten. Fehlende Zeitreihen wurden mit Hilfe der Datenbank Main Economic Indicators ergänzt. Das in den Wechselkursmodellen benötigte inländische Volkseinkommen wurde durch den Index der Industrieproduktion ersetzt, da die Erhebung der Daten des Brutto-Inlandsprodukts nur quartalsweise erfolgt. Die kürzeren Daten für Niederlande und Deutschland resultieren aus deren Teilnahme an der europäischen Währungsunion und in Folge dessen durch die Einstellung der Datenberechnung einiger Zeitreihen. Die nachfolgenden Graphiken bilden die Differenzen zwischen der jeweiligen inländischen und amerikanischen Zeitreihen ab. Ausnahme hiervon sind der Wechselkurs und die Handelsbilanz. Die starken Abweichungen nach Juli 1994 erklären die Abweichungen der ADF Testergebnisse im Vergleich zu Verkooyien.



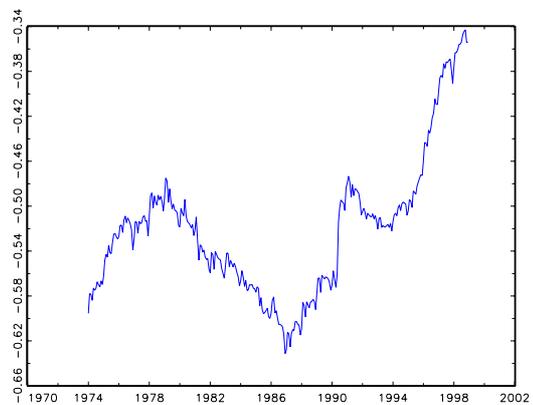
(a) Logarithmierter Wechselkurs



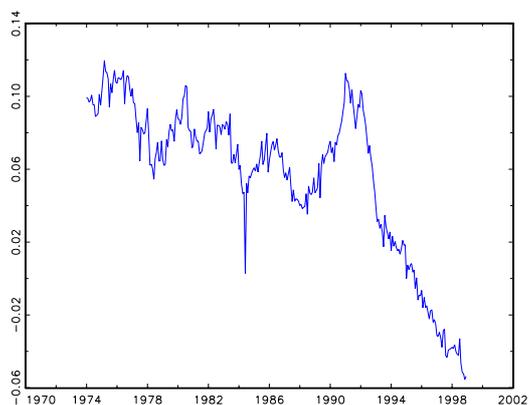
(b) Inflation



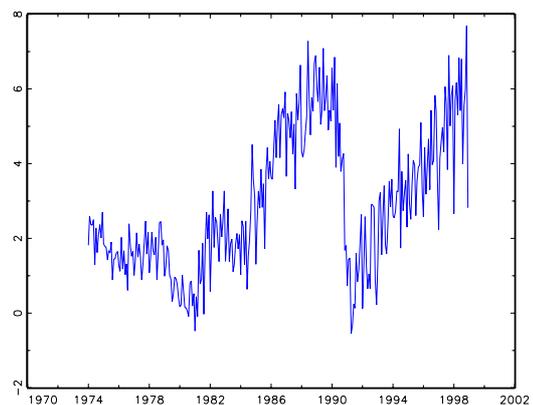
(c) Zinsdifferenz



(d) Logarithmierte Geldmengendifferenz

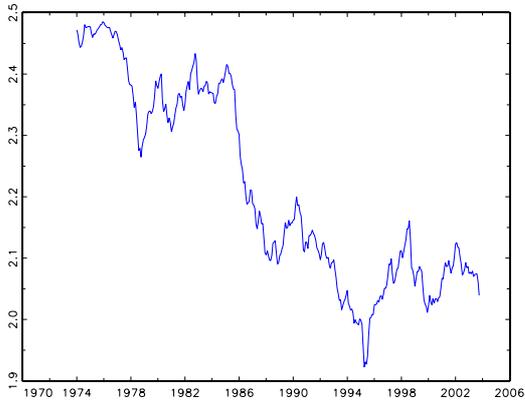


(e) Differenz der logarithmierten Industrie Produktion

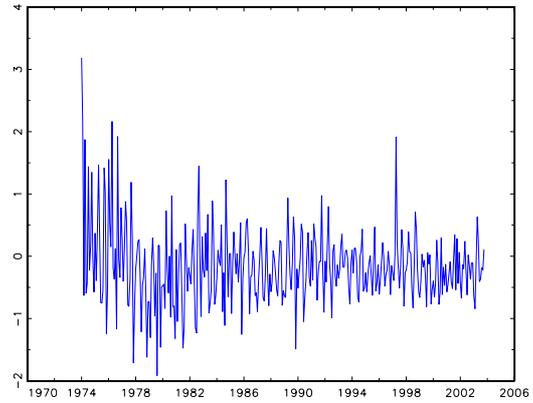


(f) Handelsbilanz BRD

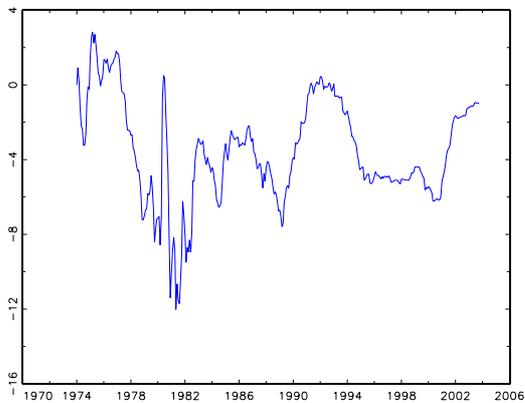
Abbildung D.1: Daten für Deutschland-USA; alle Daten bis auf Wechselkurs und Handelsbilanz sind differenziert



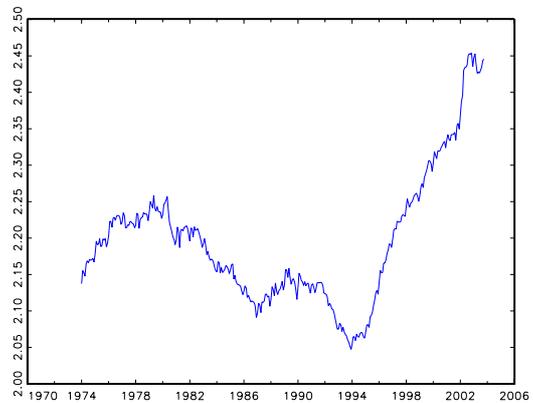
(a) Logarithmierter Wechselkurs



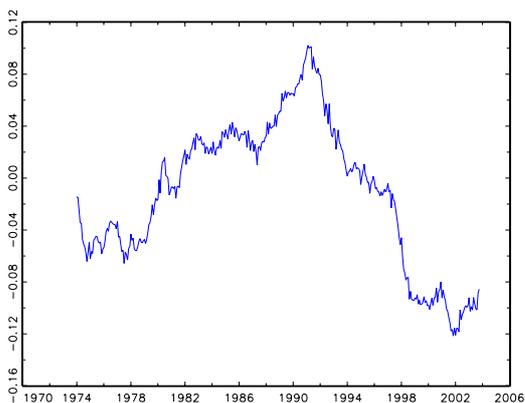
(b) Inflation



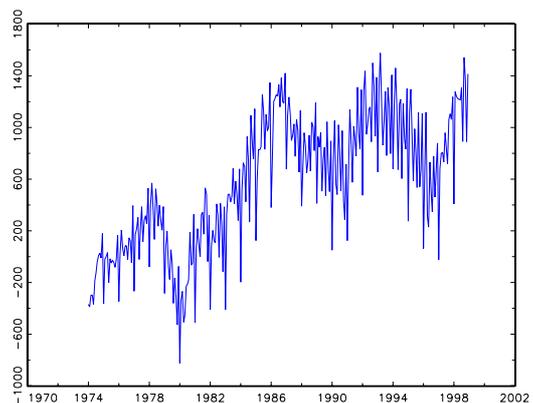
(c) Zinsdifferenz



(d) Logarithmierte Geldmengendifferenz

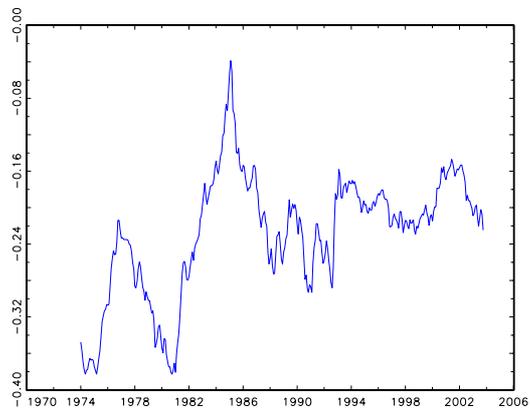


(e) Differenz der logarithmierten Industrie Produktion

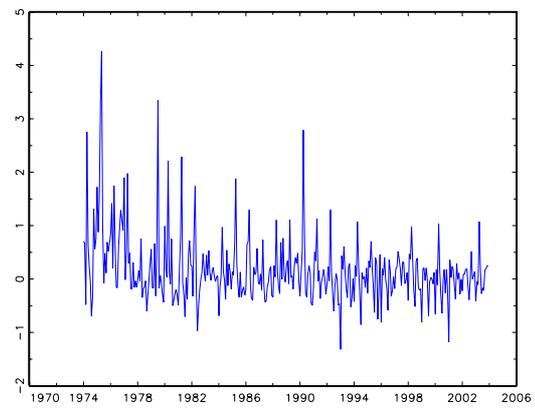


(f) Handelsbilanz Japan

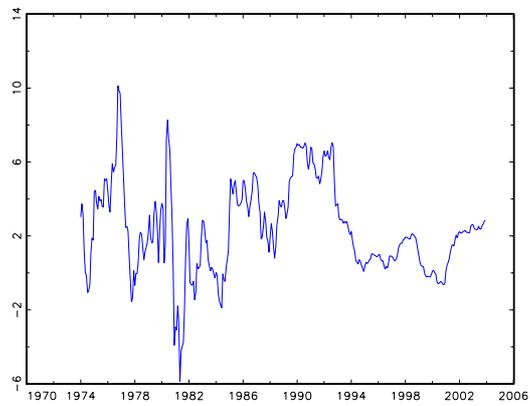
Abbildung D.2: Daten für Japan-USA; alle Daten bis auf Wechselkurs und Handelsbilanz sind differenziert



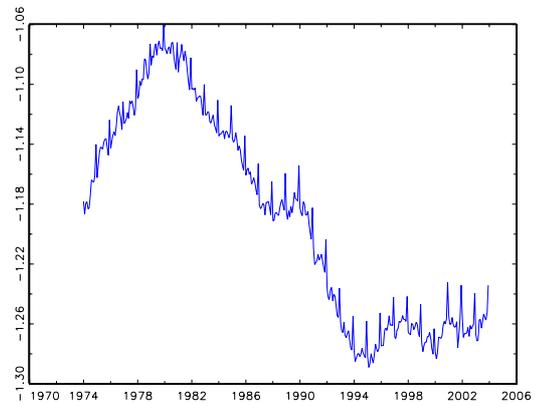
(a) Logarithmierter Wechselkurs



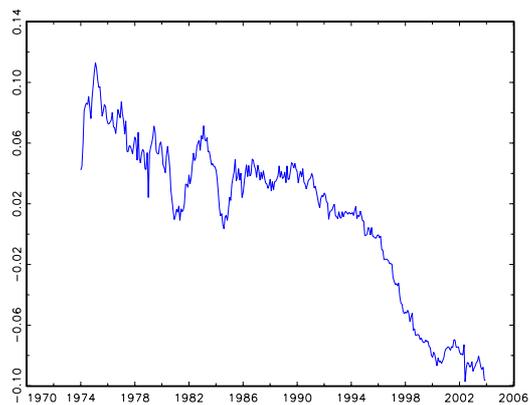
(b) Inflation



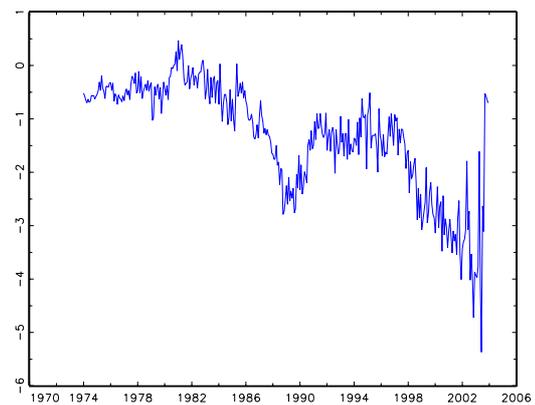
(c) Zinsdifferenz



(d) Logarithmierte Geldmengendifferenz

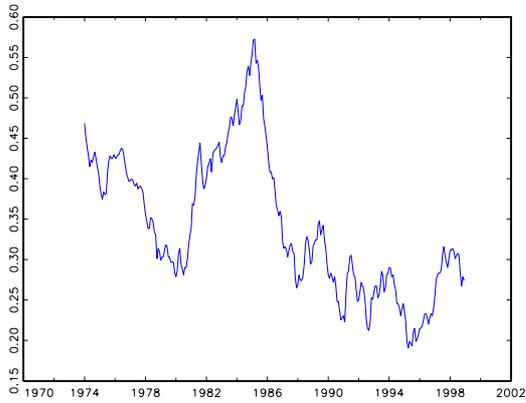


(e) Differenz der logarithmierten Industrie Produktion

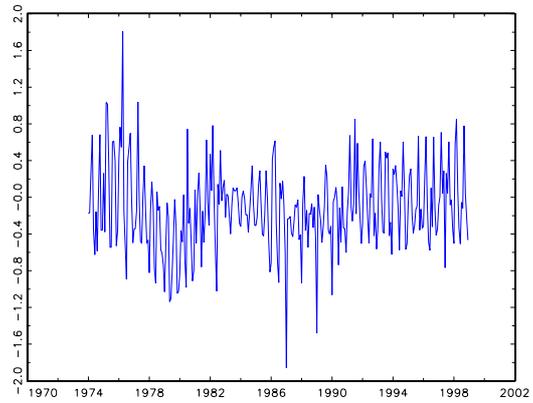


(f) Handelsbilanz GB

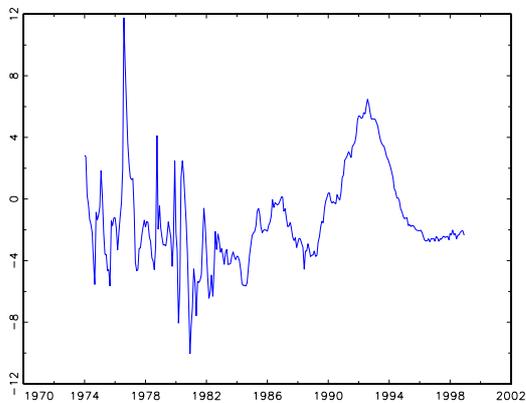
Abbildung D.3: Daten für Großbritannien-USA; alle Daten bis auf Wechselkurs und Handelsbilanz sind differenziert



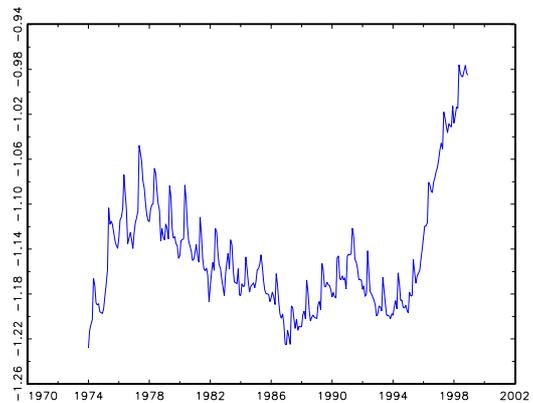
(a) Logarithmierter Wechselkurs



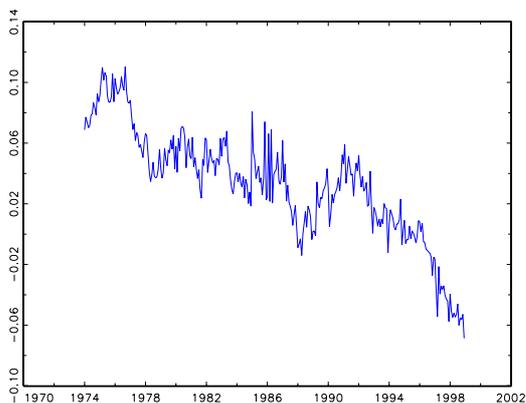
(b) Inflation



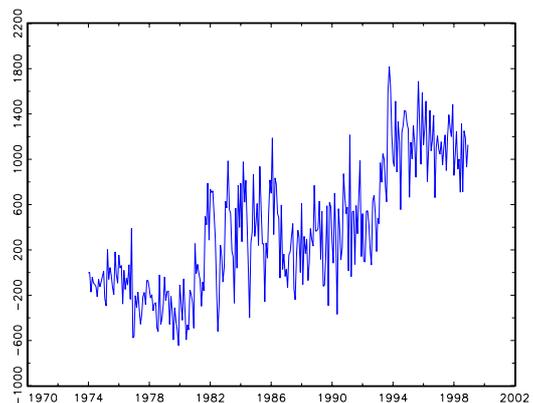
(c) Zinsdifferenz



(d) Logarithmierte Geldmengendifferenz



(e) Differenz der logarithmierten Industrie Produktion



(f) Handelsbilanz NL

Abbildung D.4: Daten für Niederlande-USA; alle Daten bis auf Wechselkurs und Handelsbilanz sind differenziert

Variable	Beschreibung	Einheit	letzter Wert	Kodierung
U.S.A.				
M1	US MONEY SUPPLY - M1 CURA	US \$ Bln	Dez.2003	USOCM1MNB
M1	US MONEY SUPPLY - M1 CURN	US \$ Bln	Dez.2003	USOCM1MNA
M0	US MONETARY BASE CURN	US \$ Bln	Dez.2003	USMONBASA
IP	US INDUSTRIAL PRODUCTION	Index	Dez.2003	USOCIPRDG
CPI	US CPI SADJ	Index	Dez.2003	USOCP009E
IR	US CERTIFICATES OF DEPOSIT	Prozent	Dez.2003	USOCTBL%
TB	US NET TRADE BALANCE	US \$ Mln	Dez.2003	USOCVBALA
Deutschland				
S	Wechselkurs - DM/U.S.-\$		Dez.1998	MEI
M1	BD MONEY SUPPLY - M1 CURA	DM Bln	Dez.1998	BDOCM1MNB
IP	BD INDUSTRIAL PRODUCTION	Index	Dez.1998	BDOCIPRDG
CPI	BD CPI NADJ	Index	Dez.1998	BDOCPCONF
IR	FIBOR 3 Monate	Prozent	Dez.1998	MEI
TB	BD NET TRADE BALANCE	DM Bln	Dez.1998	BDOCVBALA
Großbritannien				
S	Wechselkurs - /U.S.-\$		Okt.2003	MEI
M0	UK MONEY SUPPLY M0 CURN	Bln	Okt.2003	UKM0....A
IP	UK INDUSTRIAL PRODUCTION	Index	Okt.2003	UKOCIPRDG
CPI	UK CPI NADJ	Index	Okt.2003	UKOCPCONF
IR	UK INTERBANK LOANS-3-MONTH	Prozent	Okt.2003	UKOCTBL%
TB	UK NET TRADE BALANCE	Bln	Okt.2003	MEI
Niederlande				
S	Wechselkurs - DFL/U.S.-\$		Dez.1998	MEI
M1	NL MONEY SUPPLY - M1 CURN	DFL Bln	Dez.1998	NLOCM1MNA
IP	NL INDUSTRIAL PRODUCTION	Index	Dez.1998	NLOCIPRDG
CPI	NL CPI NADJ	Index	Dez.1998	NLOCPCONF
IR	NLD Official discount rate	Prozent	Dez.1998	MEI
TB	NL NET TRADE BALANCE CURN	DFL Bln	Dez.1998	NLOCVBALA
Japan				
S	Wechselkurs - Yen/U.S.-\$		Okt.2003	MEI
M1	JP MONEY SUPPLY - M1 CURA	Yen Bln	Okt.2003	JPOCM1MNB
IP	JP INDUSTRIAL PRODUCTION	Index	Okt.2003	JPOCIPRDG
CPI	JP CPI NADJ	Index	Okt.2003	JPOCPCONF
IR	JP GOVERN.SEC. YIELD-60 DAY	Prozent	Okt.2003	JPOCTBL%
TB	JP NET TRADE BALANCE CURN	Yen Mln	Okt.2003	NLOCVBALA

Tabelle D.1: Anmerkung: MEI = Main Economic Indicator (OECD); die anderen Kodierungen stellen die Datastream Kürzel dar.

D.2 GP GAUSS: langfristige Wechselkursprognosen

D.2.1 MSE Fitneßkriterium: Punkt- und Richtungsprognosen

Modell	1 Monat	6 Monate	12 Monate	18 Monate	24 Monate
Japan					
Flexible-Price	0.0184	0.170	0.290	0.989	1.842
Sticky-Price	0.0306	0.106	0.292	0.267*	0.299*
Portfolio	0.0294	0.102	0.152	0.290	1.666
Random Walk	0.0297	0.089	0.115	0.147	0.167
Großbritannien					
Flexible-Price	0.0181	0.0542	0.4720*	5292649.1*	6.4108824e+010**
Sticky-Price	0.0193	557643.12	574292.84**	592528.88***	1.5324990e+008***
Portfolio	0.0174	0.0518*	0.0748*	0.0912***	4.9488215e+056***
Random Walk	0.0167	0.0381	0.0572	0.0719	0.0826
Deutschland					
Flexible-Price	0.0261	0.0962**	0.1322**	0.1716**	0.4128**
Sticky-Price	0.0238	0.0691	0.0969**	0.1403**	0.1787**
Portfolio	0.0244	0.0808	0.1215	0.1752	0.1860
Random Walk	0.0224	0.0642	0.0984	0.1340	0.1654
Niederlande					
Flexible-Price	0.023	0.106	61213*	0.254*	0.311
Sticky-Price	0.025	0.216	0.558	0.617	0.445
Portfolio	0.023	0.068	0.109	5841	0.20
Random Walk	0.022	0.065	0.101	0.138	0.170

Tabelle D.2: Langfristige Prognosemodelle - MSE; *ein ∞ Wert (**zwei Werte) (***)drei Werte) wurde durch einen naiven Wert ersetzt. In-Sample: 1987:3 bis 1993:12. Out-of-Sample: 1994:1 bis 1998:12 bzw. 2003:10.

Modell	1 Monat	6 Monate	12 Monate	18 Monate	24 Monate
Japan					
Flexible-Price	89.09 (33.74)	83.65 (23.57)	82.65 (21.11)	84.78 (22.30)	80.23 (15.67)
Sticky-Price	70.00 (8.77)	62.50 (3.09)	58.16 (1.20)	64.13* (3.56)	51.16* (0.01)
Portfolio	69.09 (8.14)	62.50 (3.20)	58.16 (1.11)	64.13 (3.46)	56.98 (0.74)
Random Walk	53.64 (0.22)	51.92 (0.04)	41.84 (1.67)	54.35 (0.24)	42.53 (1.15)
Großbritannien					
Flexible-Price	86.36 (29.32)	62.50 (3.25)	66.33* (5.21)	57.61* (1.08)	52.33** (0.094)
Sticky-Price	84.55 (26.33)	64.42 (4.33)	61.22** (2.46)	54.35*** (0.35)	56.98*** (0.86)
Portfolio	86.36 (29.17)	66.35 (5.61)	59.18 (1.65)	58.70 (1.44)	62.79 (2.81)
Random Walk	51.82 (0.07)	50.00 (0.00)	48.98 (0.02)	54.35 (0.36)	58.62 (1.30)
Deutschland					
Flexible-Price	84.00 (11.64)	57.78** (0.45)	56.41** (0.09)	72.73** (3.55)	59.26** (0.53)
Sticky-Price	80.00 (8.98)	66.67 (2.37)	56.41** (0.26)	57.58** (0.46)	62.96** (0.95)
Portfolio	82.00 (10.24)	66.67 (2.56)	53.85 (0.10)	48.48 (0.00)	37.04 (0.95)
Random Walk	60.00 (0.96)	60.00 (0.82)	53.85 (0.05)	39.39 (0.98)	28.57 (2.63)
Niederlande					
Flexible-Price	80.00 (9.04)	53.33 (0.09)	53.85* (0.13)	51.52* (0.02)	59.25 (0.45)
Sticky-Price	84.00 (11.64)	46.66 (0.11)	61.54 (0.90)	69.70 (2.61)	59.26 (0.53)
Portfolio	86.00 (13.17)	68.88 (3.17)	56.41 (0.23)	69.70 (2.61)	40.74 (0.44)
Random Walk	56.00 (0.35)	60.00 (0.87)	53.85 (0.082)	39.39 (0.82)	32.14 (1.71)

Tabelle D.3: langfristige Prognosemodelle; Richtungsprognosen in %; **MSE als Fitneßkriterium**; In-Sample: 1987:3 bis 1993:12. Out-of-Sample: 1994:1 bis 1998:12 bzw. 2003:10. *Ein ∞ Wert (**zwei Werte usw.) wurde durch naiven Wert ersetzt. Signifikanzniveaus: 10%, 5% und 1% entsprechen Teststatistiken (in Klammern) von 2.71, 3.84 und 6.63.

D.2.2 AIC Fitneßkriterium: Punkt- und Richtungsprognosen

Modell	1 Monat	6 Monate	12 Monate	18 Monate	24 Monate
Japan					
Flexible-Price	0.033*	0.100*	0.195*	0.391*	0.461*
Sticky-Price	0.02955	0.093	0.884	0.869	0.911
Portfolio	0.031	1653671.2	1703045.4	6.3089441e+034	6.6859387e+035
Random Walk	0.02965	0.089	0.115	0.147	0.167
Großbritannien					
Flexible-Price	0.0188	0.0511*	0.0881	0.2865	1.5821
Sticky-Price	0.0170	0.0508	0.0872	0.1391	0.2069
Portfolio	0.0168	0.0367	0.0579	0.0739	0.0846
Random Walk	0.0167	0.0381	0.0572	0.0719	0.0826
Deutschland					
Flexible-Price	0.0238	0.0724	0.1006	0.1377	0.1743
Sticky-Price	0.0265	0.0735	0.1151	0.1327	0.1603
Portfolio	0.0225	0.0630	3.7070	17741.6	96076102
Random Walk	0.0224	0.0642	0.0984	0.1340	0.1654
Niederlande					
Flexible-Price	0.0224	0.0712	0.1520	0.2899	0.3221
Sticky-Price	0.1493	> 1 Mrd.	> 1 Mrd.	> 1 Mrd.	> 1 Mrd.
Portfolio	0.0240	0.0698	0.1115*	0.1615*	0.1979*
Random Walk	0.0223	0.0646	0.1010	0.1380	0.1703

Tabelle D.4: langfristige Prognosemodelle mit **AIC als Fitneßkriterium**; *Ein ∞ Wert (**zwei Werte) (***)drei Werte) wurde durch einen naiven Wert ersetzt. In-Sample: 1987:3 bis 1993:12. Out-of-Sample: 1994:1 bis 1998:12 bzw. 2003:10.

Modell	1 Monat	6 Monate	12 Monate	18 Monate	24 Monate
Japan					
Flexible-Price	68.18*	66.35*	58.16*	72.83*	59.30*
	(7.23)	(5.66)	(1.20)	(9.51)	(1.51)
Sticky-Price	70.00	70.12	62.24	66.30	62.79
	(8.69)	(8.22)	(2.58)	(4.78)	(2.71)
Portfolio	69.09	72.12	55.10	63.04	61.63
	(7.87)	(10.19)	(0.52)	(3.07)	(2.05)
Random Walk	53.64	51.92	41.84	54.35	42.53
	(0.22)	(0.04)	(1.67)	(0.24)	(1.15)
Großbritannien					
Flexible-Price	83.64	62.50*	68.37	60.87	51.16
	(24.92)	(3.25)	(6.65)	(2.17)	(0.02)
Sticky-Price	82.73	62.50	63.27	56.52	63.95
	(23.76)	(3.25)	(3.44)	(0.77)	(3.35)
Portfolio	74.55	66.35	67.35	66.30	61.63
	(13.44)	(5.56)	(5.89)	(4.89)	(2.33)
Random Walk	51.82	50.00	48.98	54.35	58.62
	(0.07)	(0.00)	(0.02)	(0.36)	(1.30)
Deutschland					
Flexible-Price	84.00	66.67	71.79	63.63	51.85
	(11.64)	(2.56)	(3.69)	(1.10)	(0.00)
Sticky-Price	84.00	71.11	58.97	63.64	51.85
	(11.56)	(3.97)	(1.00)	(1.22)	(0.03)
Portfolio	84.00	64.44	61.54	51.52	70.37
	(11.56)	(1.99)	(1.38)	(0.01)	(2.31)
Random Walk	60.00	60.00	53.85	39.39	28.57
	(0.96)	(0.82)	(0.05)	(0.98)	(2.63)
Niederlande					
Flexible-Price	88.00	53.33	53.85	57.58	51.85
	(14.56)	(0.11)	(0.19)	(0.39)	(0.00)
Sticky-Price	84.00	55.56	58.97	51.51	40.74
	(11.64)	(0.26)	(0.56)	(0.00)	(0.54)
Portfolio	84.00	68.89	56.41*	57.58*	55.55*
	(11.56)	(3.25)	(0.31)	(0.39)	(0.18)
Random Walk	56.00	60.00	53.85	39.39	32.14
	(0.35)	(0.87)	(0.082)	(0.82)	(1.71)

Tabelle D.5: langfristige Prognosemodelle; Richtungsprognosen in %; **Fitneßkriterium:** AIC; In-Sample: 1987:3 bis 1993:12. Out-of-Sample: 1994:1 bis 1998:12 bzw. 2003:10. P-Wert: 10%, 5%, 1% entsprechen Teststatistiken (in Klammern) von 2.71, 3.84, 6.63.

D.3 GP GAUSS: kurzfristige Wechselkursprognosen

D.3.1 AIC vs. MSE - Punktprognose

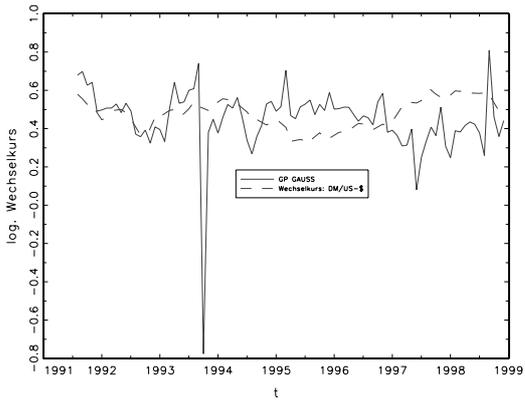
Modell	JP	GB	BRD	NL
AIC ab Louvre Abkommen				
Flexible-Price	0.0336 (0.410)	0.0189 (0.330)	0.0241 (0.295)	0.0233 (0.36)
Sticky-Price	0.0298 (0.352)	0.0176* (0.292)	0.0234 (0.328)	0.0483 —
Portfolio	0.0319* (0.321)	0.0296 —	0.0285 —	0.0334 —
Random Walk	0.0337	0.0226	0.0254	0.0245
MSE ab Louvre Abkommen				
Flexible-Price	0.0297 (0.353)	0.0183 (0.279)	0.0245 (0.289)	0.0264 —
Sticky-Price	0.0293 (0.428)	0.0172 (0.297)	0.0239 (0.306)	0.0225 (0.33)
Portfolio	0.0675 —	0.1696 —	0.0307 —	0.0298 —
Random Walk	0.0337	0.0226	0.0254	0.0245

Tabelle D.6: kurzfristige Prognosemodelle; In-Sample: 1987:3 bis 1993:12. Out-of-Sample: 1994:1 bis 1998:12 bzw. 2003:10. Punktprognosen mit Signifikanzniveaus in Klammern; *Ein ∞ Wert (**zwei Werte) (***)drei Werte) wurde durch naiven Wert ersetzt.

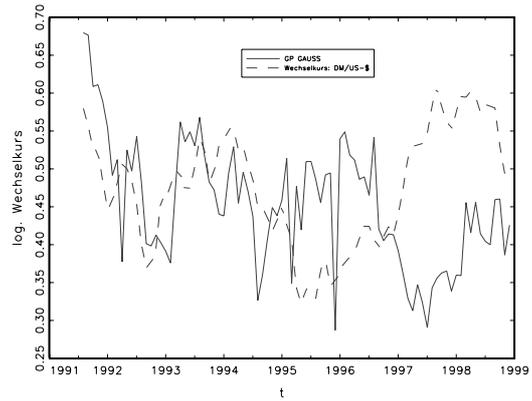
D.3.2 AIC vs. MSE - Richtungsprognose

Modell	JP	GB	BRD	NL
AIC ab Louvre Abkommen				
Flexible-Price	87.96 (31.16)	78.80 (18.36)	71.43 (4.54)	69.39 (3.72)
Sticky-Price	87.04 (29.69)	76.85* (15.88)	75.51 (6.38)	65.30 (2.17)
Portfolio	69.44* (8.20)	65.74 (5.42)	67.35 (2.98)	65.30 (2.30)
Random Walk	31.48 (7.49)	40.74 (2.00)	48.98 (0.01)	53.06 (0.03)
MSE ab Louvre Abkommen				
Flexible-Price	87.96 (31.15)	78.70 (17.96)	73.47 (5.41)	69.39 (3.72)
Sticky-Price	88.10 (48.86)	77.78 (17.09)	73.47 (5.41)	67.34 (2.89)
Portfolio	71.30 (9.79)	64.81 (40.70)	63.27 (1.72)	57.14 (0.57)
Random Walk	31.48 (7.49)	40.74 (2.00)	48.98 (0.01)	53.06 (0.03)

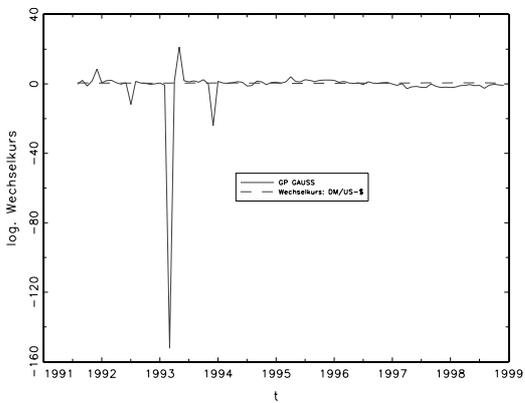
Tabelle D.7: kurzfristige Prognosemodelle; In-Sample: 1987:3 bis 1993:12. Out-of-Sample Zeitraum: 1994:1 bis 1998:12 bzw. 2003:10. Richtungsprognosen in Prozent; *Ein ∞ Wert (**zwei Werte) (***)drei Werte) wurde durch einen naiven Wert ersetzt. Signifikanzniveaus von 10%, 5% und 1% entsprechen Teststatistiken (in Klammern) von 2.71, 3.84 und 6.63.



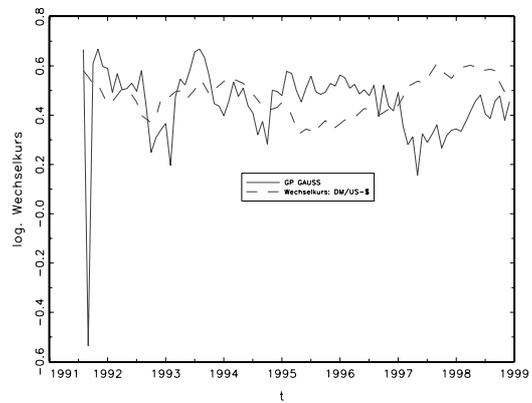
(a) Monetäre Theorie bei flexiblen Preisen: DM/US-\$ Prognose mit MSE als Fitneßkriterium



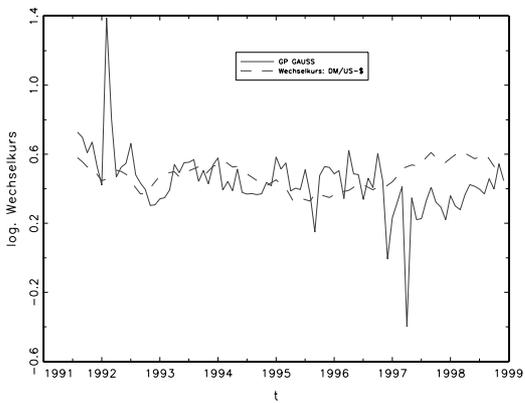
(b) Monetäre Theorie bei flexiblen Preisen: DM/US-\$ Prognose mit AIC als Fitneßkriterium



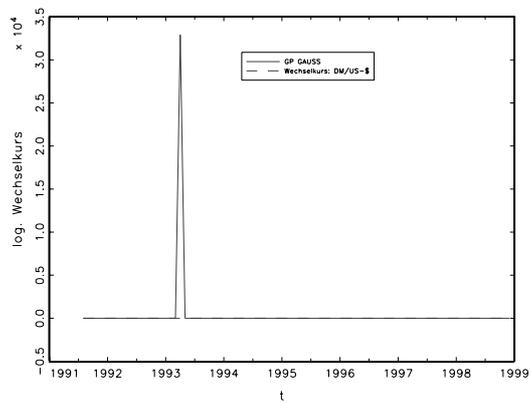
(c) Monetäre Theorie bei trägen Preisen: DM/US-\$ Prognose mit MSE als Fitneßkriterium



(d) Monetäre Theorie bei trägen Preisen: DM/US-\$ Prognose mit AIC als Fitneßkriterium

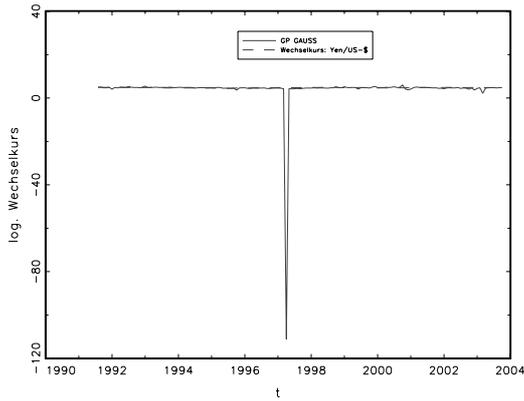


(e) Portfoliotheorie: DM/US-\$ Prognose mit MSE als Fitneßkriterium

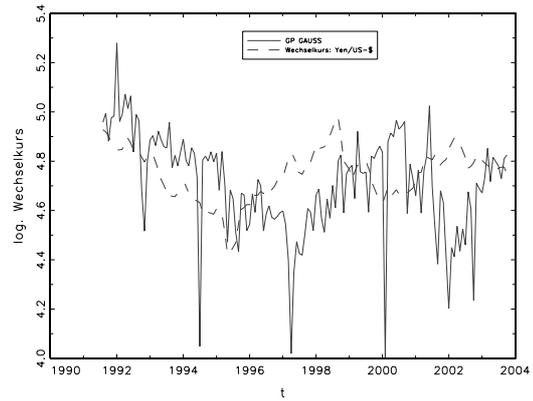


(f) Portfoliotheorie: DM/US-\$ Prognose mit AIC als Fitneßkriterium

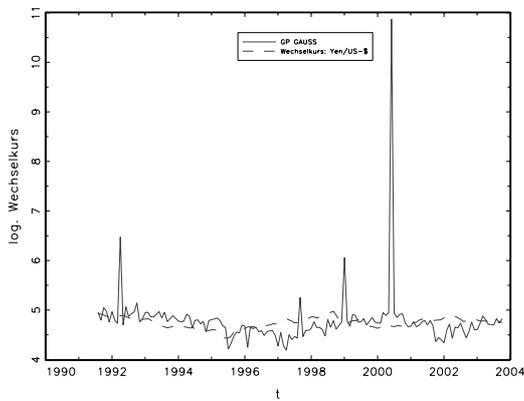
Abbildung D.5: Wechselkursvorhersagen für DM/US-\$: Prognosehorizont 24 Monaten



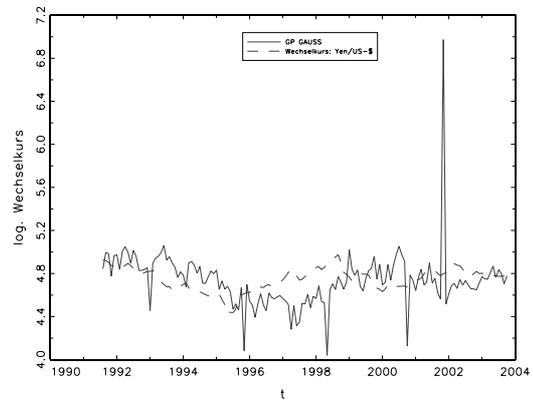
(a) Monetäre Theorie bei flexiblen Preisen: Yen/US-\$ Prognose mit MSE als Fitneßkriterium



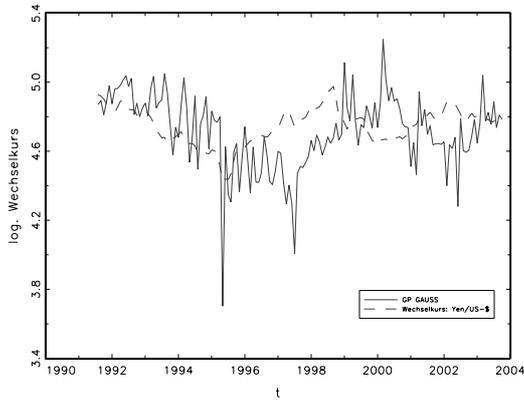
(b) Monetäre Theorie bei flexiblen Preisen: Yen/US-\$ Prognose mit MSE als Fitneßkriterium



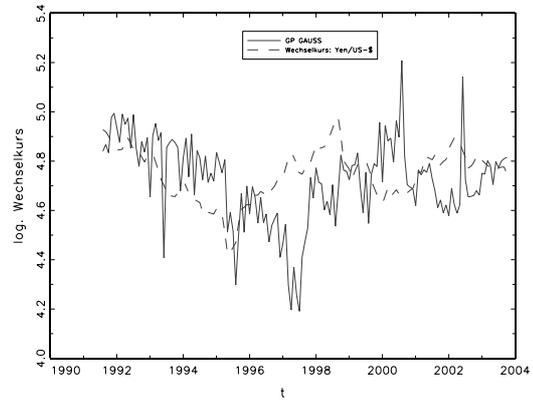
(c) Monetäre Theorie bei trägen Preisen: Yen/US-\$ Prognose mit MSE als Fitneßkriterium



(d) Monetäre Theorie bei trägen Preisen: Yen/US-\$ Prognose mit AIC als Fitneßkriterium

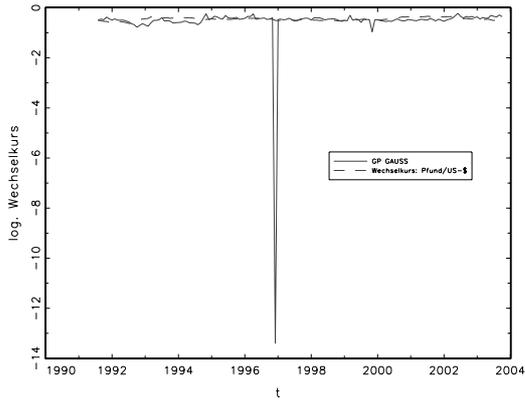


(e) Portfoliotheorie: Yen/US-\$ Prognose mit MSE als Fitneßkriterium

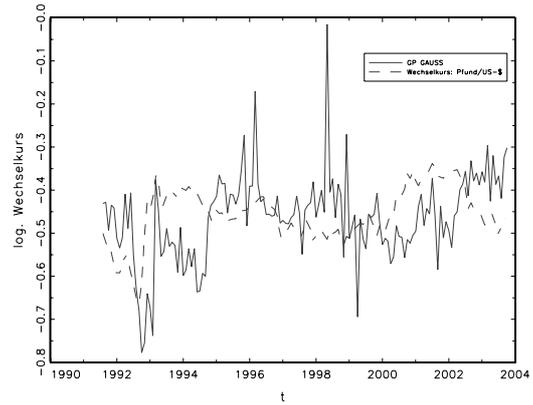


(f) Portfoliotheorie: Yen/US-\$ Prognose mit AIC als Fitneßkriterium

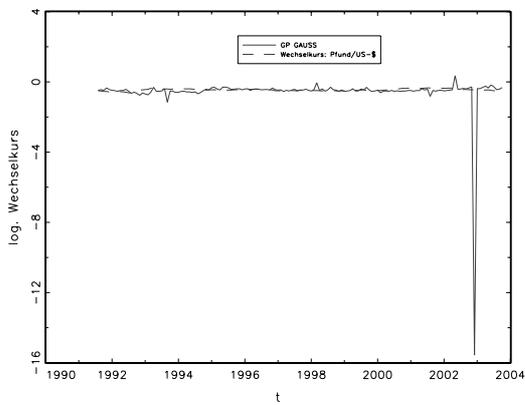
Abbildung D.6: Wechselkursvorhersagen für Yen/US-\$: Prognosehorizont 24 Monaten



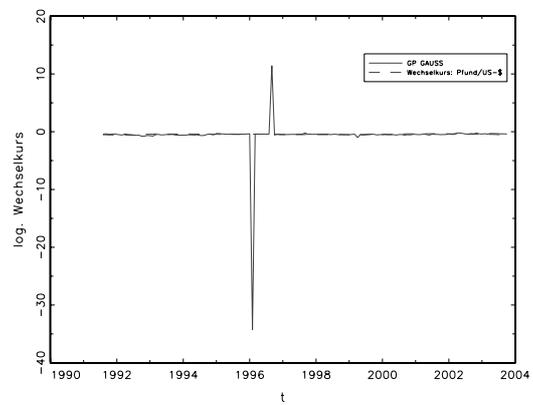
(a) Monetäre Theorie bei flexiblen Preisen: Pfund/US-\$ Prognose mit MSE als Fitneßkriterium



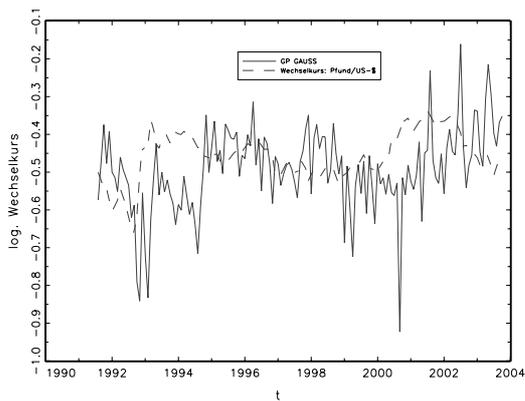
(b) Monetäre Theorie bei flexiblen Preisen: Pfund/US-\$ Prognose mit AIC als Fitneßkriterium



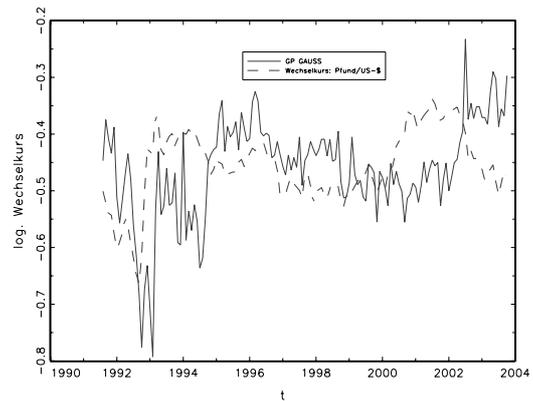
(c) Monetäre Theorie bei trägen Preisen: Pfund/US-\$ Prognose mit MSE als Fitneßkriterium



(d) Monetäre Theorie bei trägen Preisen: Pfund/US-\$ Prognose mit AIC als Fitneßkriterium

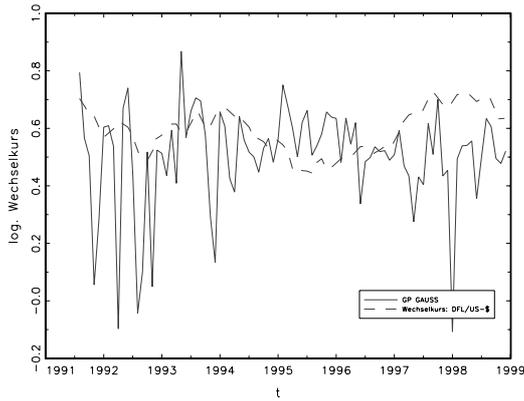


(e) Portfoliotheorie: Pfund/US-\$ Prognose mit MSE als Fitneßkriterium

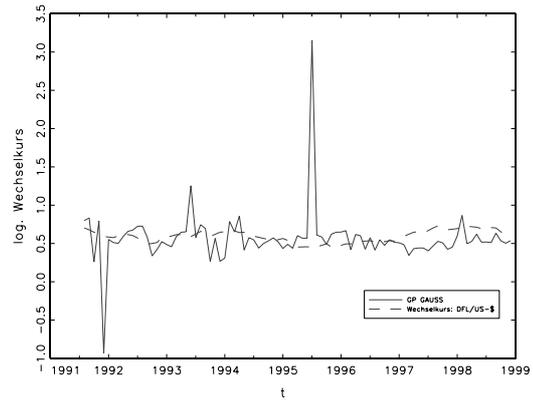


(f) Portfoliotheorie: Pfund/US-\$ Prognose mit AIC als Fitneßkriterium

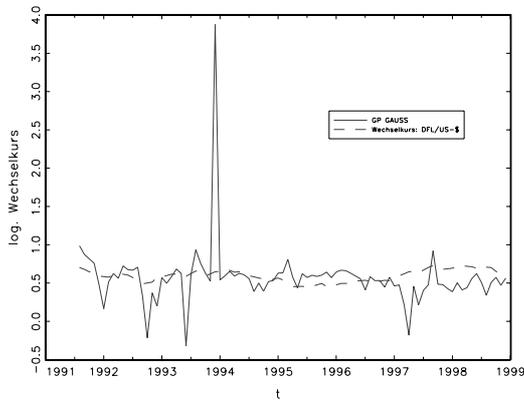
Abbildung D.7: Wechselkursvorhersagen für Pfund/US-\$: Prognosehorizont 24 Monaten



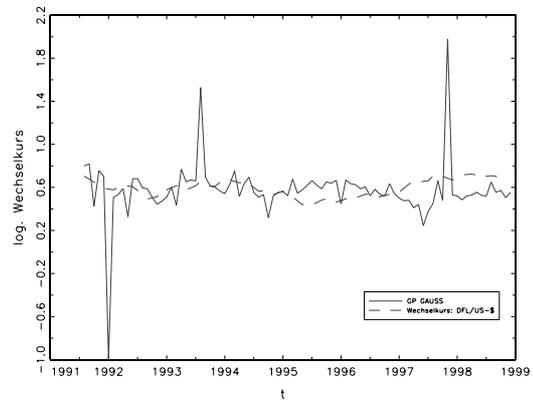
(a) Monetäre Theorie bei flexiblen Preisen: GFL/US-\$ Prognose mit MSE als Fitneßkriterium



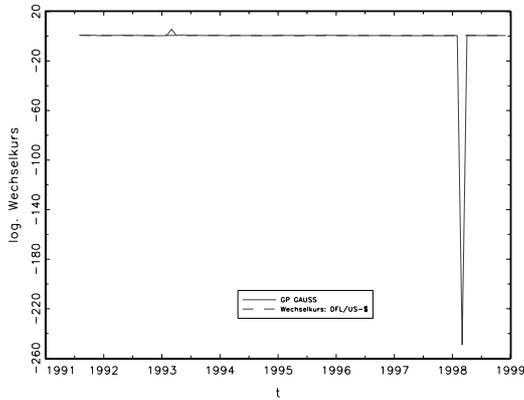
(b) Monetäre Theorie bei flexiblen Preisen: GFL/US-\$ Prognose mit AIC als Fitneßkriterium



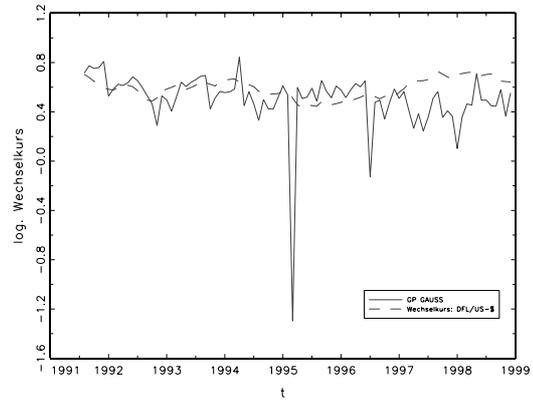
(c) Monetäre Theorie bei trägen Preisen: GFL/US-\$ Prognose mit MSE als Fitneßkriterium



(d) Monetäre Theorie bei trägen Preisen: GFL/US-\$ Prognose mit AIC als Fitneßkriterium



(e) Portfoliotheorie: GFL/US-\$ Prognose mit MSE als Fitneßkriterium



(f) Portfoliotheorie: GFL/US-\$ Prognose mit AIC als Fitneßkriterium

Abbildung D.8: Wechselkursvorhersagen für GFL/US-\$: Prognosehorizont 24 Monaten

D.4 Langfristige Wechselkursprognosen: Auszug aus der Studie von Verkooijen et al.

Modell	1 Monat	6 Monate	12 Monate	18 Monate	24 Monate
Japan					
Flexible-Price	0.09	0.15	0.22	0.28	0.34
Sticky-Price	0.10	0.19	0.25	0.32	0.38
Portfolio	0.09	0.17	0.24	0.33	0.38
Random Walk	0.03	0.08	0.11	0.14	0.17
Großbritannien					
Flexible-Price	0.11	0.15	0.17	0.19	0.24
Sticky-Price	0.15	0.13	0.17	0.20	0.21
Portfolio	0.12	0.15	0.16	0.17	0.21
Random Walk	0.04	0.13	0.17	0.19	0.20
Deutschland					
Flexible-Price	0.25	0.38	0.49	0.56	0.57
Sticky-Price	0.18	0.27	0.36	0.41	0.44
Portfolio	0.18	0.27	0.36	0.43	0.45
Random Walk	0.04	0.12	0.16	0.20	0.20
Niederlande					
Flexible-Price	0.21	0.28	0.34	0.38	0.41
Sticky-Price	0.16	0.24	0.34	0.40	0.43
Portfolio	0.13	0.12	0.23	0.27	0.29
Random Walk	0.04	0.11	0.15	0.19	0.19

Tabelle D.8: Quelle: Verkooijen et al. [127], S.23. Prognosewerte auf der Grundlage eines künstlichen neuronalen Netzwerkes.

Danksagung

An dieser Stelle gilt mein ganz besonders herzlicher Dank Professor Dr. Thomas Lux, der mir bereits bei meiner Diplomarbeit an der Universität in Bonn unterstützend zur Seite stand. Seine kritischen und konstruktiven Hinweise und Ideen haben dazu beigetragen, die Dissertation zu einem erfolgreichen Abschluß zu bringen. Nicht zuletzt wegen seiner großen Menschlichkeit war es mir eine Freude, als Assistent am Lehrstuhl zu arbeiten.

Ich bedanke mich auch bei meinen Arbeitskollegen Eleni Samanidou, Florian Heitger, Hwataek Lee und Simone Alfarano für die konstruktiven Diskussionen aber vor allem für die harmonische gemeinsame Arbeitszeit.

Weiterhin möchte ich mich bei Thomas Hahn-Bruckart sowie meinem Bruder Ansgar Zschischang bedanken, die viel Mühe und Zeit für das Korrekturlesen investiert haben. Mein besonderer Dank gilt auch meinen Eltern für ihre motivierende und in jeder Hinsicht aufbauende Unterstützung.

Kiel, April 2005

Literaturverzeichnis

- [1] M. Adya and F. Collopy. How effective are Neural Networks at forecasting and prediction? A review and evaluation. *Journal of Forecasting*, 17:481–495, 1998.
- [2] P.J. Angeline. Genetic Programming and emergent intelligence. In K.E. Kinnear, Jr., editor, *Advances in Genetic Programming 2*, pages 89–110. MIT Press, Cambridge, MA., 1994.
- [3] R. Ashley and D.M. Patterson. *A Nonlinear Time Series Workshop*. Kluwer Academic Publishers, Norwell, 2000.
- [4] R. Baille and P. McMahon. *The Foreign Exchange Market: Theory and Evidence*. Cambridge University Press, Cambridge, 1989.
- [5] W. Banzhaf, P. Nordin, R.E. Keller, and F.D. Francone. *Genetic Programming*. Morgan Kaufmann, San Francisco, 1998.
- [6] C. Barnhart. *Nonlinearity and Business Cycles*. Unpublished dissertation, Clemson University, 1995.
- [7] J. Beran. *Statistics for Long-Memory Processes*. Chapman and Hall, New York, 1994.
- [8] J.F. Bilson. Recent developments in monetary models of exchange rate determination. *IMF Staff Papers*, 26 (2):201–223, 1979.
- [9] J. Binner, G. Kendall, and A. Gazely. Evolving Neural Networks with Evolutionary Strategies: A new application to divisia money. In Thomas B. Fomby, editor, *Advances in Econometrics*. Elsevier Science Ltd., Oxford, 2004.
- [10] M. Blix, J. Wadefjord, U. Wienecke, and M. Adahl. How good is the forecasting performance of major institutions? *Economic Review*, 3:38–68, 2001.
- [11] P. Bofinger, J. Reischle, and A. Schächter. *Geldpolitik - Ziele, Institutionen, Strategien und Instrumente*. Franz Vahlen, München, 1996.
- [12] J.-P. Bouchaud and M. Potters. *Theory of Financial Risk and Derivative Pricing: From Statistical Physics to Risk Management*. Cambridge University Press, Cambridge, 2003.

- [13] G.E.P. Box and G.M. Jenkins. *Time Series Analysis. Forecastin and Control*. Holden-Day, San Francisco, 1975.
- [14] W.A. Brock, W.D. Dechert, and J. Scheinkman. A test for independence based on the correlation dimension. *Workshop Papers SSRI-8702*, 1987. University of Wisconsin-Madison.
- [15] W.A. Brock, D.A. Hsieh, and B. LeBaron. *Nonlinear Dynamics, Chaos, and Instability: Statistical Theory and Economic Evidence*. MIT Press, Cambridge, 1991.
- [16] P.J. Brockwell and R.A. Davis. *Time Series: Theory and Methods*. Springer, New York, 1991.
- [17] N. Cameron and S. Moshiri. Econometrics vs. ANN models in forecasting inflation. *Journal of Forecasting*, 19:201–217, 2000.
- [18] K. Carstensen and J. Hawallek. Forecasting inflation from the term structure. *Arbeitspapier, Christian-Albrechts-Universität zu Kiel*, 154:1–15, 2001.
- [19] B. Chen and D.M. Titterington. Neural Networks: A review from a statistical perspective. *Statistical Science*, 1:2–54, 1994.
- [20] S.-H. Chen. *Evolutionary Computation in Economics and Finance*. Physica-Verlag, Heidelberg, 2002.
- [21] S.-H. Chen. *Genetic Algorithms and Genetic Programming in Computational Finance*. Kluwer Academic Publishers, Massachusetts, 2002.
- [22] S.-H. Chen. *Computational Intelligence in Economics and Finance*. Springer-Verlag, Heidelberg, 2003.
- [23] Y. Cheung, M. Chinn, and A. Pascual. Empirical exchange rate models of the nineties:are any fit to survive? University of California Santa Cruz, 2002. NBER Working Paper.
- [24] N. Chidambaran, C.-W. J. Lee, and J. Trigueros. Option pricing via Genetic Progamming. In S.-H. Chen, editor, *Evolutionary Computation in Economics and Finance*, pages 383–398. Physica-Verlag, 2002.
- [25] C. Darwin. *On the Origin of Species by Means of Natural Selection or the Preservation of Favoured Races in the Struggle for Life*. Murray, London, 1859.
- [26] R.S. Didner. Intelligent systems at american express. In S. und P. Treleven Gooatilake, editor, *Intelligent Systems for Finance and Business*, pages 31–37. Wiley, Chichester, 1995.
- [27] F. Diebold and R. Mariano. Comparing predicitive accuracy. *Journal of Business and Economic Statistics*, 13:253–263, 1995.

- [28] F.X. Diebold. *Elements of Forecasting*. South Western College Publishing, Cincinnati, 2000.
- [29] F.X. Diebold and J.A. Nason. Nonparametric exchange rate prediction? *Journal of International Economics*, 28:315–332, 1990.
- [30] R. Dornbusch. Expectations and exchange rate dynamics. *Journal of Political Economy*, 84:1161–1176, 1976.
- [31] A.E. Drake and R.E. Marks. Genetic Algorithms in economics and finance: Forecasting stock market and foreign exchange - a review. In S.-H. Chen, editor, *Genetic Algorithms and Genetic Programming in Computational Finance*, pages 29–54. Kluwer Academic Publishers, Massachusetts, 2002.
- [32] D. Duwendag, K.-H. Ketterer, W. Kösters, R. Pohl, and D.B. Simmert. *Geldtheorie und Geldpolitik in Europa*. Springer, Heidelberg, 1999.
- [33] R.F. Engle. Autoregressive conditional heteroskedasticity with estimates of the variance of United Kingdom inflation. *Econometrica*, 50:987–1007, 1982.
- [34] R.F. Engle and C.W.J. Granger. Cointegration and error-correction: Representation, estimation and testing. *Econometrica*, 55:251–276, 1987.
- [35] A. Episcopos and J. Davis. Predicting returns on Canadian exchange rates with Artificial Neural Networks and EGARCH-M models. *Neural Computing and Applications*, 4:168–174, 1996.
- [36] A. Estrella and Mishkin F. The predictive power of the term structure of interest rates in Europe and the United States: Implications for the European Central Bank. *European Economic Review*, 41:1375–1401, 1997.
- [37] E.F. Fama. Short-term interest rates as predictors of inflation. *American Economic Review*, 65:269 ff., 1975.
- [38] E.F. Fama. The information in the term structure. *Journal of Financial Economics*, 13(4):509–528, 1984.
- [39] E.F. Fama. Term-structure forecasts of interest rates, inflation and real return. *Journal of Monetary Economics*, 25:59–60, 1990.
- [40] E.F. Fama and M.R. Gibbons. Inflation, real returns and capital investment. *Journal of Monetary Economics*, 9:297–324, 1982.
- [41] I. Fisher. *The Theory of Interest*. MacMillan, New York, 1930.
- [42] D. Fogel. *Evolutionary Computation*. IEEE PRESS Marketing, New York, 1995.
- [43] L. Fogel, A. Owens, and M. Walsh. Artificial intelligence through a simulation of evolution. In Callahan A. und L. Fogel Maxfield, M., editor, *Biophysics and Cybernetic Systems*, pages 131–155. 1965.

- [44] J. Frankel. On the mark: A theory of floating exchange rates based on real interest differentials. *The American Economic Review*, 69:611–622, 1979.
- [45] J.A. Frenkel. A monetary approach to the exchange rate: Doctrinal aspects and empirical evidence. *Scandinavian Journal of Economics*, 78:200–224, 1976.
- [46] M. Gärtner. *Macroeconomics under Flexible Exchange Rates*. Harvester Wheatsheaf, New York, 1995.
- [47] C. Gathercole and P. Ross. Small populations over many generations can beat large populations over few generations in Genetic Programming. *Genetic Programming 1997: Proceedings of the Second Annual Conference*, pages 111–118, 1997.
- [48] J. Geweke and S. Porter-Hudak. The estimation and application of long-memory time series models. *Journal of Time Series Analysis*, 4:221–238, 1983.
- [49] Gil-Aluja. *Fuzzy Sets in the Management of Uncertainty*. Springer-Verlag, Heidelberg, 2004.
- [50] D.E. Goldberg and K. Sastry. Modeling tournament selection with replacement using apparent added noise. In L. Spector, E.D. Goodman, A. Wu, W.B. Langdon, H.-M. Voigt, M. Gen, S. Sen, M. Dorigo, S. Pezeshk, M.H. Garzon, and E. Burke, editors, *Proceedings of the Genetic and Evolutionary Computation Conference (GECCO-2001)*, page 781 ff. Morgan Kaufmann, San Francisco, 7-11 2001.
- [51] S. Goonatilake and P. Treleaven. *Intelligent Systems for Finance and Business*. Wiley, Chichester, 1995.
- [52] C. Granger and A. Andersen. *An Introduction to Bilinear Time Series Models*. Vandenhoeck and Ruprecht, Gottingen, 1978.
- [53] C.W.J. Granger and R. Joyeux. An introduction to long-memory models and fractional differencing. *Journal of Time Series Analysis*, 1:15–39, 1980.
- [54] C.W.J. Granger and P. Newbold. Spurious regressions in econometrics. *Journal of Econometrics*, 35:143–159, 1974.
- [55] D. Guillaume. *Intradaily Exchange Rate Movements*. Boston, 2000.
- [56] U. Haßler. *Fraktional integrierte Prozesse in der Ökonometrie*. Dissertation, Institut für Statistik und Ökonometrie, Freie Universität Berlin, 1993.
- [57] G. Hansen. Prognostiziert die Zinsstruktur die Inflation in Deutschland? *Kredit und Kapital*, 4:554–578, 2001.
- [58] H. Hansen and K. Juselius. Cats in rats: Cointegration analysis of time series, 1995. Manuscript.
- [59] C. Harff. *Indexierte Staatsanleihen und Geldpolitik*. Dissertation, Institut für Volkswirtschaftslehre, Christian-Albrechts-Universität zu Kiel, 2000.

- [60] D.I. Harvey, S.J. Leybourne, and P. Newbold. Testing the equality of prediction mean squared errors. *International Journal of Forecasting*, 13:281–291, 1997.
- [61] M. Henon. A two-dimensional mapping with a strange attractor. *Comm. Math. Phys.*, 50:69, 1976.
- [62] M. Hinich. Testing for Gaussianity and linearity of a stationary time series. *Journal of Time Series Analysis*, 3:169–176, 1982.
- [63] M. Hinich. Testing for dependence in the input to a linear time series model. *Non-parametric Statistics*, 6:205–221, 1996.
- [64] J. Holland. *Adaption in Natural and Artificial Systems*. MIT Press, Cambridge, MA, 1975.
- [65] J.R.M. Hosking. Fractional differencing. *Biometrika*, 68:165–176, 1981.
- [66] D.A. Hsieh. Testing for nonlinear dependence in daily foreign exchange rates. *Journal of Business*, 62:329–368, 1989.
- [67] H.E. Hurst. Long-term storage of capacity of reservoirs. *Transactions of the American Society of Civil Engineers*, 116:770–799, 1951.
- [68] G. Illing. *Theorie der Geldpolitik - Eine spieltheoretische Einführung*. Springer, 1997.
- [69] C.M. Jarque and A.K. Bera. A test for normality of observations and regression residuals. *International Statistical Review*, 55:163–172, 1987.
- [70] Ch. Jochum and G. Kirchgässner. Hat die Zinsstruktur Aussagekraft für die zukünftige Inflation in Deutschland? Eine Kritik des Mishkin-Ansatzes. *Kredit und Kapital*, 4:493–519, 1999.
- [71] S. Johansen. Statistical analysis of cointegrated vectors. *Journal of Economic Dynamics and Control*, 12:231–254, 1988.
- [72] M. Kaboudan. Statistical properties of fitted residuals from genetically evolved models. *Journal of Economic Dynamics and Control*, 25:1719–1749, 2001.
- [73] M. Kaboudan and M. Vance. Statistical evaluation of symbolic regression forecasting of time-series. University of Cambridge, 1998. Proceedings of the intentional federation of automatic control symposium on computation in economics, finance and engineering: economic systems.
- [74] L. Kanzler. Very fast and correctly sized estimation of the bds statistic. <http://ssrn.com/abstract=151669>, 1999. Working Paper.
- [75] J.R. Koza. *Genetic Programming*. MIT Press, Cambridge, 1992.

- [76] U. Kretschmer. *Der Informationsgehalt von Inflationserwartungen der Finanzmärkte für die Inflationsprognose*. Diplomarbeit, Institut für Statistik und Ökonometrie, Universität Kiel, 2000.
- [77] C.-M. Kuan and T. Liu. Forecasting exchange rates using feedforward and recurrent Neural Networks. *Journal of Applied Econometrics*, 10, No.4:347–364, 1995.
- [78] C.-T. Lin. *Neural Fuzzy Control Systems With Structure and Parameter Learning*. World Scientific Publishing Company, 1994.
- [79] T. Liu, C. Granger, and W. Heller. Using the correlation exponent to detect whether an economic series is chaotic. *Journal of Applied Econometrics*, 7:25–29, 1992.
- [80] G.M. Ljung and G.E.P. Box. On a measure of lack of fit in time series models. *Biometrika*, 65:297–303, 1978.
- [81] A.W. Lo. Long-term memory in stock markets prices. *Econometrica*, 59:1279–1313, 1991.
- [82] H. Lütkepohl. *Introduction to Multiple Time Series Analysis*. Springer Verlag, 1991.
- [83] M. Mackey and L. Glass. Oscillations and chaos in physiological control systems. *Science*, 197:287, 1977.
- [84] J.G. MacKinnon. Critical values of cointegration tests. In R.F. und C.W.J. Granger Engle, editor, *Long-Run Econometric Relationships: Readings in Cointegration*. Oxford University Press, New York, 1991.
- [85] B.B. Mandelbrot and J.W. Ness. Fractional brownian motions, fractional noises and applications. *SIAM Review*, 10:422–437, 1968.
- [86] N. Mark. Exchange rates and fundamentals: Evidence on long horizon predictability. *American Economic Review*, 85:201–218, 1995.
- [87] R. May. Simple mathematical models with very complicated dynamics. *Nature*, 261:459–467, 1976.
- [88] A.I. McLeod and W.K. Li. Diagnostic checking ARMA time series models using squared-residual autocorrelations. *Journal of Time Series Analysis*, 4:269–273, 1983.
- [89] R. Meese and K. Rogoff. Empirical exchange rate models of the seventies: Do they fit out of sample? *Journal of International Economics*, 14:3–24, 1983.
- [90] F. Mishkin. The real rate of interest: An empirical investigation. *The Cost and Consequences of Inflation - Carnegie Rochester Conference Series on Public Policy*, 15:151–200, 1981.
- [91] F. Mishkin. The information in the longer maturity term structure about future inflation. *The Quarterly Journal of Economics*, 105:815–828, 1990.

- [92] F. Mishkin. What does the term structure tell us about future inflation? *Journal of Monetary Economics*, 25:77–95, 1990.
- [93] F. Mishkin. Is the Fisher effect for real? A reexamination of the relationship between inflation and interest rates. *Journal of Monetary Economics*, 30:195–215, 1992.
- [94] T. Mitze. *Forecasting the EMU Inflation Rate. A Fuzzy Rule-Based Approach with GENEFER Compared to Mainstream Econometrics*. PhD thesis, Institute for Computational Economics, University of Münster, 2002.
- [95] C.R. Nelson, G. Schwert, and G. William. Short-term interest rates as predictors of inflation: On testing the hypothesis that the real rate of interest is constant. *American Economic Review*, 67:478–86, 1977.
- [96] C.R. Nelson and A.F. Siegel. Parsimonious modeling of yield curves. *Journal of Business*, 60,4:473–489, 1987.
- [97] P. Nordin and W. Banzhaf. Complexity compression and evolution. In L. Eshelman, editor, *Genetic Algorithms: Proceedings of the Sixth International Conference (ICGA95)*, pages 318–325. Morgan Kaufmann, San Francisco, CA., 1995.
- [98] P. Nordin, F. Francone, and W. Banzhaf. Explicitly defined introns and destructive crossover in Genetic Programming. In Rosca, editor, *Proceedings of the Workshop on Genetic Programming: From Theory to Real World Applications*, pages 6–22. Tahoe City, CA., 1995.
- [99] U. O’Reilly. *An Analysis of Genetic Programming*. Phd thesis, Carleton University, Ottawa, Kanada, 1995.
- [100] T. Ozaki. Non-linear threshold autoregressive models for non-linear random vibrations. *Journal of Applied Probability*, 18:443–451, 1981.
- [101] K. Patterson. *An Introduction to Applied Econometrics*. Palgrave, 2000.
- [102] W. Pedrycz. *Fuzzy Evolutionary Computation*. Kluwer Academic Publishers, 1997.
- [103] K. Pilbeam. *International Finance*. Palgrave Macmillan, London, 1992.
- [104] J. Plasmans, W. Verkooijen, and H. Daniels. Estimating structural exchange rate models by Artificial Neural Networks. *Applied Financial Economics*, 8:541–551, 1998.
- [105] R. Poli and W.B. Langdon. Schema theory for Genetic Programming with one-point crossover and point mutation. *Evolutionary Computation*, 6(3):231–252, 1998.
- [106] F.A. Rauscher. Exchange rate forecasting: A neural VEC approach. *Neural Network World: International Journal on Neural and Mass-Parallel Computing and Information Systems*, 7:461–471, 1997.

- [107] I. Rechenberg. *Cybernetic Solution Path of an Experimental Problem*. Farnborough, Hants., UK, 1965.
- [108] A.-P. N. Refenes, Y. Abu-Mostafa, J. Moody, and A. Weigend. *Neural Networks in Financial Engineering*. World Scientific, Singapore, 1996.
- [109] R.A. Ribeiro, H.-J. Zimmermann, R.R. Yager, and J. Kacprzyk. *Soft Computing in Financial Engineering*. Physica-Verlag, Heidelberg, 1999.
- [110] R. Richter. Notenbank und Wahrung in Deutschland seit 1948. *Deutsche Bundesbank - Funfzig Jahre Deutsche Mark*, 1998.
- [111] J.P. Rosca. Analysis of complexity drift in Genetic Programming. In Koza et al., editor, *Genetic Programming 1997: Proceedings of Second Annual Conference*, pages 286–294. Morgan Kaufmann, Stanford University, 1997.
- [112] A.L. Samuel. Some studies in machine learning using the game of checkers. *IBM Journal*, 3:211–229, 1959.
- [113] M. Schroder. *Finanzmarktokometrie*. Schaffer-Poeschel, Stuttgart, 2002.
- [114] H.-P. Schwefel. *Kybernetische Evolution als Strategie der experimentellen Forschung in der Stromungstechnik*. Diplomarbeit, Technische Universitat Berlin, 1965.
- [115] H.R. Seddighi, K.A. Lawler, and A.V. Katos. *Econometrics: A Practical Approach*. Routledge, London, 2000.
- [116] T. Soule, J.A. Foster, and J. Dickinson. Code growth in Genetic Programming. In J.R. Koza, D.E. Goldberg, D.B. Fogel, and R.L. Riolo, editors, *Genetic Programming 1996: Proceedings of the First Annual Conference*, pages 215–223. MIT Press, Stanford University, 1996.
- [117] L. Svensson. Inflation forecast targeting: Implementing and monitoring inflation targets. *European Economic Review* 41, 6:1111–1146, 1997.
- [118] L.E.O. Svensson. Estimating and interpreting forward interest rates: Sweden 1992-94, 1994. IWF Working Paper 114.
- [119] N.R. Swanson and H. White. A model selection approach to real-time macroeconomic forecasting using linear models and Artificial Neural Networks. *The Review of Economics and Statistics*, LXXIX, No.1:540550, 1997.
- [120] H. Takagi. Introduction to Fuzzy Systems, Neural Networks, and Genetic Algorithms. In D. Ruan, editor, *Intelligent Systems: Fuzzy, Logic, Neural Networks, and Genetic Algorithms*. Norwell, 1997.
- [121] W.A. Takett. *Recombination, Selection, and the Genetic Constuction of Computer Programs*. Phd thesis, Department of Electrical Engineering Systems, University of Southern California, 1994.

- [122] A. Tettamanzi and M. Tomassini. *Soft Computing: Integrating Evolutionary, Neural, and Fuzzy Systems*. Springer, Heidelberg, 2001.
- [123] H. Tong. *Nonlinear Time Series Analysis: A Dynamical System Approach*. PhD thesis, Oxford University Press, Oxford, 1990.
- [124] R.R. Trippi and E. Turban. *Neural Networks in Finance and Investing*. Irwin, Chicago, 1996.
- [125] R.S. Tsay. Nonlinearity tests for time series. *Biometrika*, 73:461–466, 1986.
- [126] E. Tzavalis and M.R. Wickens. Forecasting inflation from the term structure. *Journal of Empirical Finance*, 3:103–122, 1996.
- [127] W. Verkooijen. A Neural Network approach to long-run exchange rate prediction. *Computational Economics*, 9:51–65, 1996.
- [128] R.F. Walker, E.W. Haasdijk, and M.C. Gerrets. Credit evaluation using a Genetic Algorithm. In S. Goonatillake and P. Treleaven, editors, *Intelligent Systems for Finance and Business*. Wiley, Chichester, 1995.
- [129] P. Whigham. *Grammatical Bias for Evolutionary Learning*. PhD thesis, University of New South Wales, Australia, 1996.
- [130] M. Willms. *Internationale Währungspolitik*. Vahlen, München, 1992.
- [131] J. Wolters. Untersuchung der Renditestruktur am deutschen Kapitalmarkt 1970-1996. In Baltensperger E., editor, *Spekulation, Preisbildung und Volatilität auf Finanz- und Devisenmärkten*, pages 129–152. Duncker und Humblot, 1998.

CURRICULUM VITAE

ELMAR ZSCHISCHANG



PERSÖNLICHE ANGABEN

Geburtsdatum: 11. September 1973
Geburtsort: Minden (Westfalen), Deutschland
Staatsbürgerschaft: Deutsch

AUSBILDUNG

- 07/1986–05/1993 Allgemeine Hochschulreife am Humboldtgynasium in Wilhelmshaven
- 10/1994–07/1996 Diplom-Vorprüfungs-Zeugnis für den Studiengang Volkswirtschaftslehre an der Rheinischen Friedrich-Wilhelms-Universität Bonn
- 08/1996–02/2000 Diplomprüfungs-Zeugnis für den Studiengang Volkswirtschaftslehre an der Rheinischen Friedrich-Wilhelms-Universität Bonn
- Titel der Diplomarbeit: *Mikrosimulationsmodelle für Finanzmärkte: Das Modell von Levy und Solomon*; Betreuer: Professor Dr. Thomas Lux
- 11/2000-11/2004 Wissenschaftlicher Angestellter an der Christian-Albrechts-Universität zu Kiel sowie Koordinator der Kiel Summer School der Wirtschafts- und Sozialwissenschaftlichen Fakultät