

Übungsserie 1: Zeitreihenanalyse

Aufgabe 1: Stromproduktion

Die Daten sind unter Zeitreihen_Übungsserie 1.xls auf Moodle zu finden.

Importieren Sie aus Excel die Zeitreihe der Stromproduktion. Tabellenblatt **Üb1Data**

gretl: Daten als Zeitreihen importieren

gretl: Tabellenkalk.-Import

Starte Import bei:

Spalte: 2 Zeile: 2

(B)

Tabellenblatt zum Import:

☒ b1Data

Struktur des Datensatzes

☐ Querschnittsdaten

☒ Zeitreihe

Zeitreihenfrequenz

☐ Jährlich

☒ Quartalsweise

Startbeobachtung

Quartalsweise 2010:1

Zuerst Datei auswählen, dann Spalte/Zeile ausfüllen

1. Erklären Sie kurz den Unterschied zwischen der Ökonometrie und Zeitreihenanalyse

Ökonometrie = äussere Methode → die Entwicklung einer Zeitreihe wird durch bestimmte erklärende Variablen erklärt, die als kausale Einflussgrössen in Fragen kommen

2. In welche Komponenten lassen sich ökonomische Zeitreihen zerlegen? Wie lassen sich die Komponenten inhaltlich erklären?

Ökonomische Zeitreihen lassen sich in **systematische** und eine **Residualkomponente** zerlegen.

Während die systematischen Komponenten gewisse Regelmässigkeiten enthalten, wird die Residualkomponente bei der Zeitreihenzerlegung als **Restkomponente** aufgefasst, die durch Zufallseinflüsse und irreguläre Bewegungen geprägt ist.

Die **systematische Komponente** umfasst die **glatte Komponente**, die aus Trend und Konjunktur (=zyklische Komponente) besteht, und bei **unterjährigen** Zeitreihen zusätzlich die **Saisonkomponente**.

Der **Trend**, der z.B. auf den technischen Fortschritt, das Bevölkerungswachstum oder Produktzyklen zurückgeführt werden kann, spiegelt die **langfristige Bewegungsrichtung** einer Zeitreihe wider.

Die **Konjunkturkomponente** gibt das Auf und Ab der wirtschaftlichen Aktivität um den Trend wieder.

Die zyklischen Konjunkturschwankungen haben eine variable Wellenlänge von etwa 4 bis 8 Jahren.

Die z.B. in Quartals- oder Monatsdaten vorzufindende **Saisonkomponente** hat dagegen eine **konstante Wellenlänge** von einem Jahr. Sie ist auf natürliche oder institutionell bedingte jahreszeitliche Einflüsse zurückzuführen. Natürliche Einflüsse sind z.B. witterungsbedingte Einflüsse, während institutionell bedingte Einflüsse z.B. durch Feiertage, Ferien, und Zahlungstermine hervorgerufen werden.

3. Nennen Sie 5 verschiedene Zwecke der Zeitreihenanalyse

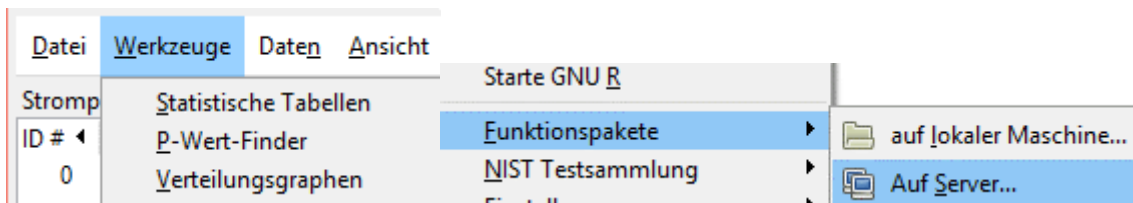
Ziele der Zeitreihenanalyse können sein

- die kürzestmögliche **Beschreibung** des historischen Verlaufs einer Zeitreihe.

- die **Vorhersage** von künftigen Zeitreihenwerten (**Prognose**) auf der Basis der Kenntnis ihrer bisherigen Werte (Umsatz- und Wettervorhersage)
- Kontrolle der zeitlichen Entwicklung → Erkennung von Veränderungen in Zeitreihen (**EEG** in der Medizin bei chirurgischen Eingriffen, Veränderung der Geldmenge in einem Land..)
- Diagnose der aktuellen Tendenz einer Zeitreihe durch die Eliminierung von seriellen oder saisonalen Abhängigkeiten (**Saisonbereinigung**) → saisonbereinigte Arbeitslosigkeit.
- Strukturerkennung durch die Identifikation des datenerzeugenden Prozesses (= wahres Modell)

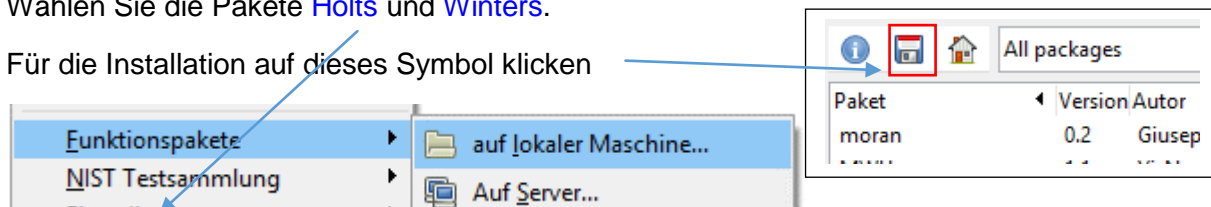
Installieren Sie die Pakete Holt und Winters auf Ihren PC.

gretl Hauptfenster: Werkzeuge / Funktionspakete / Auf Server / Namen suchen

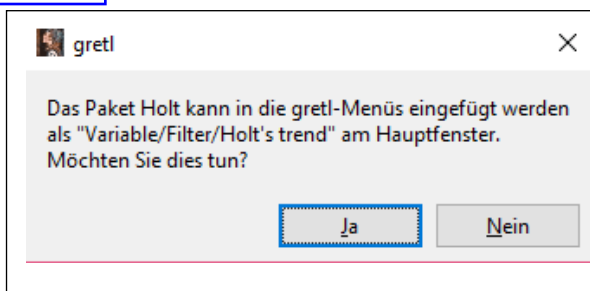


Wählen Sie die Pakete **Holts** und **Winters**.

Für die Installation auf dieses Symbol klicken



Holt	1.0	Ignacio Diaz-Emparanza	Holt's local linear trend
Winters	1.1	Ignacio Diaz-Emparanza	Smoothing and forecasting method for seasonal...



Kontrollieren, dass beide Pakete auf Ihrem PC installiert sind.



Sie erhalten die **vierteljährliche Stromproduktion** eines Kraftwerks und wollen die Zeitreihe mit unterschiedlichen Verfahren **glätten**.

4. **Regressieren Sie die Stromproduktion als Funktion der Zeit.**

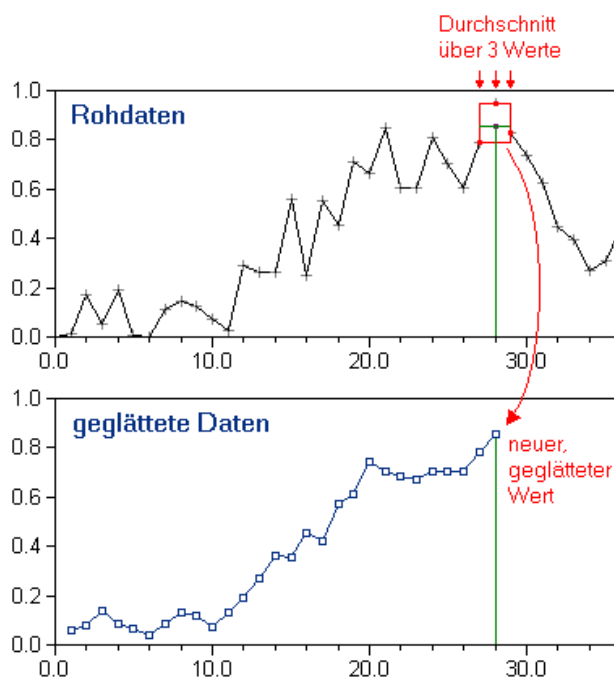
Abhängige Variable: Strom

	Koeffizient	Std.-fehler	t-Quotient	p-Wert
const	96,6973	3,80622	25,41	1,50e-015 ***
time	4,74354	0,317737	14,93	1,40e-011 ***
Mittel d. abh. Var.	146,5045	Stdabw. d. abh. Var.	29,17438	
Summe d. quad. Res.	1208,452	Stdfehler d. Regress.	8,193669	
R-Quadrat	0,925274	Korrigiertes R-Quadrat	0,921122	
F(1, 18)	222,8795	P-Wert (F)	1,40e-11	

5. Interpretieren Sie den Steigungskoeffizienten.

Die Stromproduktion steigt durchschnittlich um 4.74 GWh pro Quartal.

6. Neben der Bestimmung von Trends mittels OLS-Schätzung ist häufig die Glättung von Zeitreihen von Bedeutung. Erklären Sie kurz was mit Glättung einer Zeitreihe gemeint ist.



Glättung bedeutet das Ausschalten von unregelmässigen Schwankungen der Zeitreihe durch lokale Approximationen wie beispielsweise das lokale arithmetische Mittel.

7. Erklären Sie kurz was ein gleitender Durchschnitt 4-ter Ordnung ist.

Der einfache gleitende Durchschnitt (*moving average* (MA)) n -ter Ordnung einer diskreten Zeitreihe $x(t)$ ist die Folge der arithmetischen Mittelwerte von n aufeinanderfolgenden Datenpunkten.

8. Berechnen Sie die gleitenden Durchschnitte 4-ter und 2-ter Ordnung.

$$\text{Formel: } m_{MA}^n(t) = \frac{1}{n} \sum_{i=0}^{n-1} x(t-i)$$

Index MA = moving average = gleitender Durchschnitt

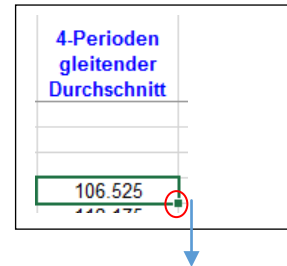
9. Berechnen Sie in Excel diese Zeitreihen. Tabellenblatt **Üb1**. (F5)

Gleitender Durchschnitt 4-ter Ordnung: F5

Gleitender Durchschnitt 2-ter Ordnung: G3

Sie können die Spalte bis Zelle 21 schnell ausfüllen,

indem Sie das **Ausfüllkästchen**  nach unten ziehen.



10. Erstellen Sie mittels gretl die Zeitreihen der gleitenden Durchschnitte 2-ter und 4-ter Ordnung.

Speichern Sie die geglätteten Reihen als **ma_Strom2** bzw. **ma_Strom4**

Variable	Modell	Hilfe
Zeige Werte	Periodogramm	
Bearbeite Attribute	Filter	Einfacher gleitender Durchschnitt
	X-12-ARIMA-Analyse	Exponentieller gleitender Durchschnitt

Einfacher gleitender Durchschnitt

Anzahl Beobachtungen im Durchschnitt:

☐ Zentriert

☒ Plote originale und geglättete Reihen

☐ Plote Rest- oder zyklische Reihe

☒ Geglättete Reihe speichern als

Einfacher gleitender Durchschnitt

Anzahl Beobachtungen im Durchschnitt:

☐ Zentriert

☒ Plote originale und geglättete Reihen

☐ Plote Rest- oder zyklische Reihe

☒ Geglättete Reihe speichern als

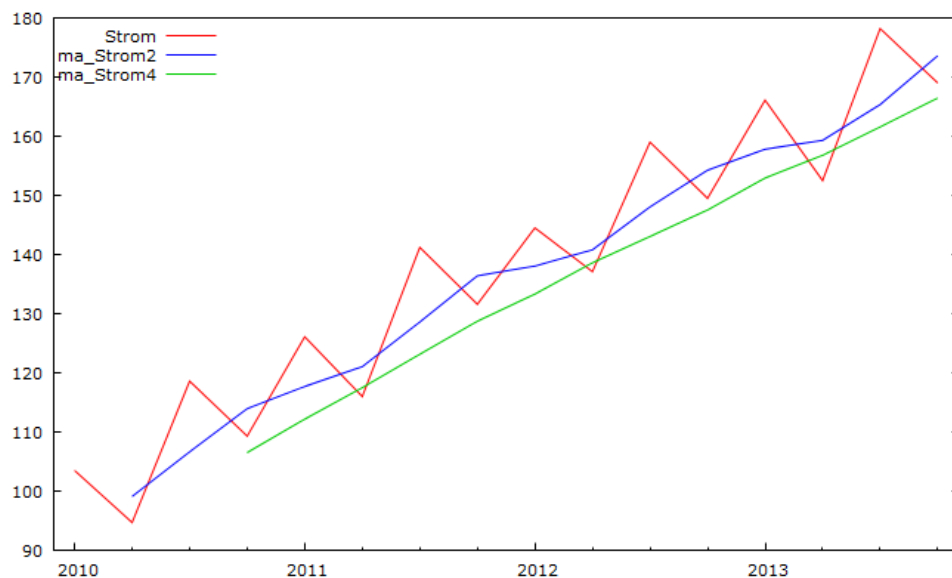
11. Zeigen Sie die originale Zeitreihe zusammen mit beiden gleitenden Durchschnitten.

gretl Hauptfenster: Ansicht / Plote spezifizierte Variablen / Zeitreihengraph

Ansicht	Hinzufügen	Stichprobe	Variable	Modell	Hilfe
Symbolansicht					
Plote spezifizierte Variablen					
			Zeitreihengraph...		

Wähle zu plottende Variablen

v1 Strom T	Strom ma_Strom2 ma_Strom4
------------------	---------------------------------



12. Welcher Unterschied ist zu vermerken?

Je höher die Ordnung der Glättung, desto stabiler die erstellte Zeitreihe.
Für die Trendbestimmung ist der 4-gliedrige gleitende Durchschnitt geeignet.

13. Glätten Sie die Zeitreihe der **Stromproduktion** in Excel mit einem **zentrierten** gleitenden Durchschnitt der 4-ten und 3-ten Ordnung! (H4)

Zentrierter 4-gliedriger gleitender Durchschnitt:

$$\bar{y}_t^4 = \frac{1}{4} (y_{t-4} + \dots + y_{t-1} + y_t + y_{t+1} + \dots + y_{t+4}) = \frac{1}{p} \sum_{k=-4}^4 y_{t+k}$$

Gleitender Durchschnitt 4-ter Ordnung: H4

Gleitender Durchschnitt 2-ter Ordnung: I3

14. Erstellen Sie mittels gretl die Zeitreihen der **zentrierten** gleitenden Durchschnitte 3-ter und 4-ter Ordnung.

Speichern Sie die geglätteten Reihen als **maz_Strom3** bzw. **maz_Strom4**

Einfacher gleitender Durchschnitt

Anzahl Beobachtungen im Durchschnitt:

☒ **Zentriert**

☒ Plote originale und geglättete Reihen

☐ Plote Rest- oder zyklische Reihe

☒ Geglättete Reihe speichern als

15. Wie viele gleitende Durchschnitte lassen sich an den **Rändern** nicht berechnen? Wann fängt der erste Wert der Zeitreihe an?

$p = 3 \rightarrow$ ungerade Zahl $\rightarrow 3 = 2q + 1 \rightarrow q = 1$ lässt sich nicht beobachten

Erster Wert: $(p+1)/2 = 2$ (Zeile)

$p = 4 \rightarrow$ gerade Zahl $\rightarrow 4 = 2k \rightarrow q = 2$ lassen sich nicht berechnen

Erster Wert: $p/2 + 1 = 2 + 1 = 3$ (Zeile)

	$p = 2q + 1$	$p = 2k$ (gerade)
Es fallen weg	$q = (p-1)/2$	$q = p/2$
Der erste Wert	$q+1 = (p+1)/2$	$q+1 = p/2 + 1$

16. Wann wird die **exponentielle Glättung** angewendet?

Die exponentielle Glättung wird vor allem verwendet, wenn die Zeitreihe keinerlei systematisches Muster wie linearen Anstieg erkennen lässt.

Das Verfahren wird beispielsweise in der Lagerhaltung verwendet, um den Bedarf eines zu bestellenden Artikels im kommenden Monat zu ermitteln.

Die exponentielle Glättung ist ein geeignetes Verfahren zur kurzfristigen Prognose aus einer Stichprobe mit periodischen Vergangenheitsdaten

17. Erklären Sie kurz das **Verfahren** der exponentiellen Glättung. Unterscheiden Sie die zwei Extremwerte $\alpha = 0$ und 1

\hat{y}_{t+1} : ermittelter Prognosewert für **t+1** zum Zeitpunkt **t** (Ein-Schritt-Prognose)


Rekursive Formel: $\hat{y}_{t+1} = \alpha y_t + (1 - \alpha) \hat{y}_t$ (mit $0 < \alpha < 1$).

Der Prognosewert für die **Periode t+1** zum Zeitpunkt **t** \hat{y}_{t+1} = gewogenes arithmetisches Mittel aus y_t , (dem tatsächlichen Wert der Periode t) und \hat{y}_t dem zur Zeit **t-1** für t prognostizierten Wert \rightarrow **Gewichtung** = Parameter α :

$\alpha = 0$: $\hat{y}_{t+1} = \hat{y}_t$ (eine einmal gestellte Prognose wird unabhängig von der Erfahrung beibehalten)

$\alpha = 1$: $\hat{y}_{t+1} = y_t$ (es wird quasi angenommen, dass morgen das eintreten wird, was heute eingetreten ist)

18. Glätten Sie die Zeitreihe der Stromproduktion in Excel mit der **exponentiellen Glättung** erster Ordnung.

Nehmen Sie $\alpha = 0.3$ in Zelle J24 = absolute Referenz $\rightarrow \$J24\$$ sonst können Sie nicht mit dem **Ausfüllkästchen**  nach unten ziehen.

Die Spalte M stellt die prognostizierten Werte dar. Spalte N berechnet die quadrierten Fehler. Die Formeln sind schon vorhanden.

19. Interpretieren Sie den Wert 176.8 in Zelle M21.

$$176.8 = 0.3 \times 195 + 0.7 \times 169.026$$

176.8 ist die Prognose für die Stromproduktion im 4. Quartal 2014 aufgrund des beobachteten Produktionswertes im 3. Quartal (195) und der im 2. Quartal prognostizierten Wert für das 3. Quartal.

20. Berechnen Sie den mittleren quadratischen Prognosefehler (N22) und die entsprechende quadratische Abweichung (RMSE) (N23).
21. Finden Sie den optimalen Glättungsparameter, indem Sie den MSE mittels Excel-Solver minimieren.

Ziel festlegen:

Bis: ☐ Max. ☒ Min. ☐ Wert:

Durch Ändern von Variablenzellen:

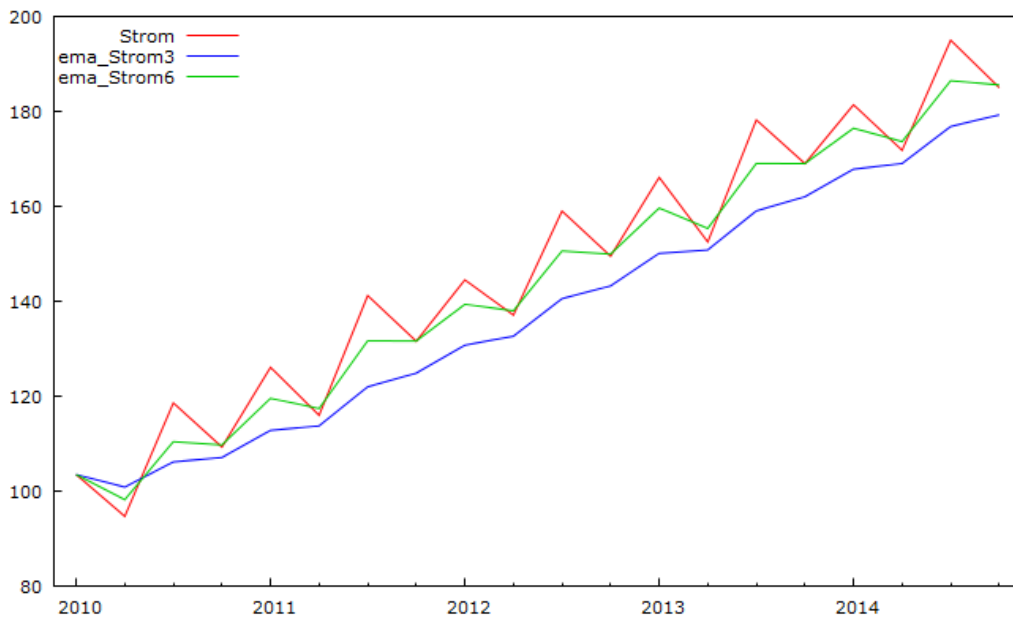
Unterliegt den Nebenbedingungen:

22. Was wäre die Prognose für das 1. Quartal 2015 anhand des optimalen Glättungsparameter α ? 185.66 (Zelle K21)
23. Erstellen Sie mittels gretl die Zeitreihen der exponentiellen Glättungen mit den Parameter $\alpha = 0.3$ und 0.6 .

Exponentieller gleitender Durchschnitt	
Gewicht auf derzeitiger Beobachtung:	<input type="text" value="0,300"/>
The initial EMA value is	
<input checked="" type="radio"/> das Mittel der ersten n Beobachtungen	<input type="text" value="1"/>
<input type="radio"/> das Mittel der Gesamtreihe	
<input type="radio"/> a user-specified value	<input type="text" value="0,1"/>
<input checked="" type="checkbox"/> Plote originale und geglättete Reihen	
<input type="checkbox"/> Plote Rest- oder zyklische Reihe	
<input checked="" type="checkbox"/> Geglättete Reihe speichern als	<input type="text" value="ema_Strom3"/>

Exponentieller gleitender Durchschnitt	
Gewicht auf derzeitiger Beobachtung:	<input type="text" value="0,600"/>
The initial EMA value is	
<input checked="" type="radio"/> das Mittel der ersten n Beobachtungen	<input type="text" value="1"/>
<input type="radio"/> das Mittel der Gesamtreihe	
<input type="radio"/> a user-specified value	<input type="text" value="0,1"/>
<input checked="" type="checkbox"/> Plote originale und geglättete Reihen	
<input type="checkbox"/> Plote Rest- oder zyklische Reihe	
<input checked="" type="checkbox"/> Geglättete Reihe speichern als	<input type="text" value="ema_Strom6"/>

24. Zeigen Sie die originäre Zeitreihe zusammen mit den beiden exponentiellen Glättungen.



25. Erstellen Sie mittels gretl die Zeitreihen der **exponentiellen Glättungen** mit den Parameter $\alpha = 0.6$. Wählen Sie für den Startwert den **Durchschnitt der ersten 4 Beobachtungen (ZR4)**

Exponentieller gleitender Durchschnitt

Gewicht auf derzeitiger Beobachtung: 0,600

The initial EMA value is

☒ das Mittel der ersten n Beobachtungen 4

☐ das Mittel der Gesamtreihe

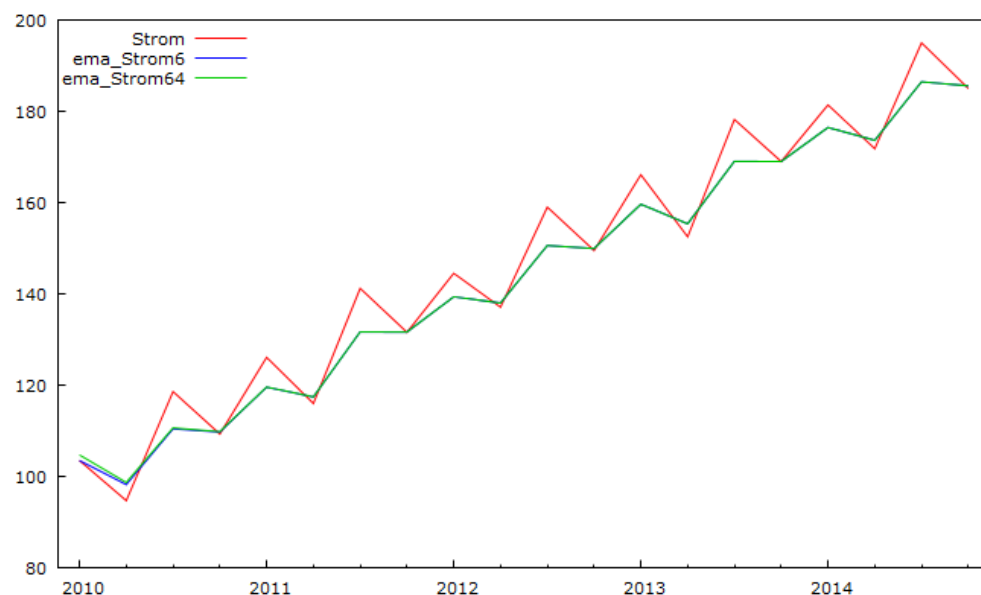
☐ a user-specified value 0,00

☒ Plote originale und geglättete Reihen

☐ Plote Rest- oder zyklische Reihe

☒ Geglättete Reihe speichern als ema_Strom64

26. Zeigen Sie die **originäre Zeitreihe** zusammen mit den beiden exponentiellen Glättungen mit $\alpha = 0.6$.



Aufgabe 2: Glättung mit dem Holt-Verfahren

Die Zeitreihe stellt die wöchentlichen Verkäufe von Thermostaten dar.
Die Daten sind in der Excel-Datei im Tabellenblatt **Üb2Data** zu finden.
Importieren Sie aus Excel die Zeitreihe der Thermostaten-Verkäufe.

gretl: Daten als Zeitreihen importieren

Zeitreihenfrequenz: Andere 1

Die genaue Wochenbezeichnung spielt bei dieser Übung keine Rolle

Zuerst Datei auswählen, dann Spalte/Zeile ausfüllen

Struktur des Datensatzes

☐ Querschnittsdaten

☒ Zeitreihe

Starte Import bei:

Spalte: 2 Zeile: 1

(B)

Tabellenblatt zum Import:

☒ b1

☒ b1_S

☒ b2Data

Zeitreihenfrequenz

☐ Dekadisch

☒ Andere 1

27. Wann wird das Holt Verfahren angewendet?

Die Zeitreihe weist einen Trend auf, dessen Steigung sich ändern kann.

Die Zeitreihe sollte keine Saisonkomponente aufweisen.

Das Glättungsverfahren erfolgt anhand von zwei Parametern α und γ .

28. Glätten Sie mittels gretl die Zeitreihe y mit den Parametern $\alpha = 0.2$, 0.6 und $\gamma = 0.1$?

Stichprobe Variable Model

Zeige Werte

Korrelogramm

Periodogramm

Filter

Fraktionale Differenz

Holt's trend

Winters

Holt's trend

Argumente auswählen:

Dependent variable (series) y

Level smoothness parameter (scalar) 0.2

Slope smoothness parameter (scalar) 0.1

Initial values based on: a regression v

Holt's trend

Argumente auswählen:

Dependent variable (series) y

Level smoothness parameter (scalar) 0,6

Slope smoothness parameter (scalar) 0,1

Initial values based on: a regression

Strom (series: original series)

Strom_Holt (series: Holt's linear smoother)

Name der Variable: y_Holt21

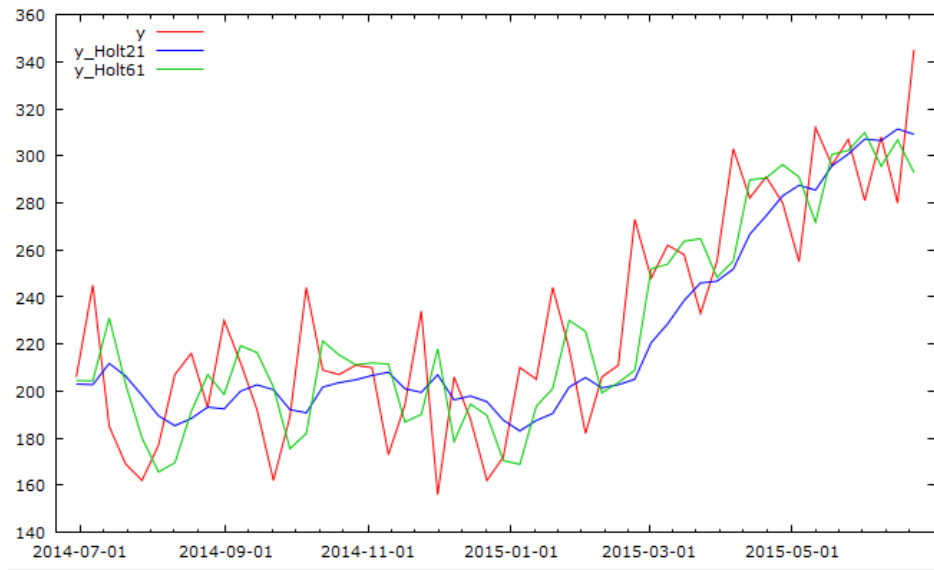
Beschreibung:

Holt's linear smoother

Speichern Sie die geglättete Zeitreihe ($\alpha = 0.2$; $\gamma=1$) als y_Holt21

Speichern Sie die geglättete Zeitreihe ($\alpha = 0.6$; $\gamma=1$) als y_Holt61

29. Zeigen Sie in einem Diagramm die ursprüngliche Zeitreihe y zusammen mit den zwei geglätteten Zeitreihen Holt21 und Holt61.

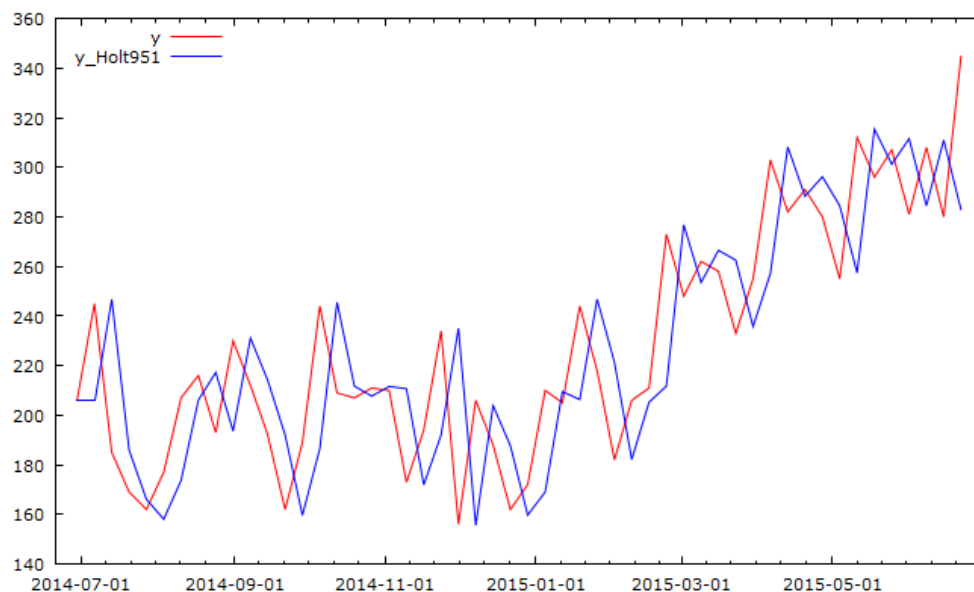


30. Was passiert wenn der Niveauparameter α steigt?

Der Glättungseffekt nimmt ab und die erstellte Zeitreihe nähert sich der ursprünglichen Zeitreihe y .

Niveauschätzung: $L_t = \alpha y_t + (1-\alpha) \hat{y}_t(t-1)$ α steigt \rightarrow 2. Term hat immer weniger Gewichtung \rightarrow Niveauschätzung reagiert schneller auf Änderungen von $y \Leftrightarrow$ die letzte Beobachtung y_t erhält eine grössere Gewichtung, je näher der Glättungsparameter α an 1 liegt.

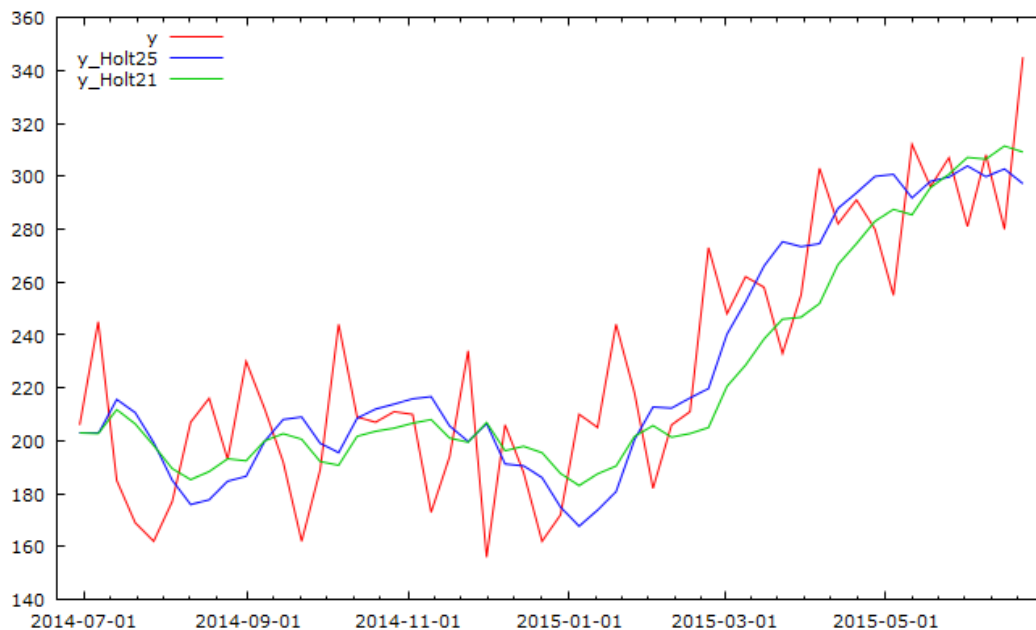
Beispiel für $\alpha = 0.95$



31. Glätten Sie mittels gretl die Zeitreihen mit den Parametern $\alpha = 0.2$ und $\gamma = 0.5$

Speichern Sie die geglättete Zeitreihe als y_Holt25

32. Zeigen Sie ein Diagramm die ursprüngliche Zeitreihe y zusammen mit den zwei geglätteten Zeitreihen Holt21 und Holt25



33. Was passiert wenn der Trendparameter γ sinkt?

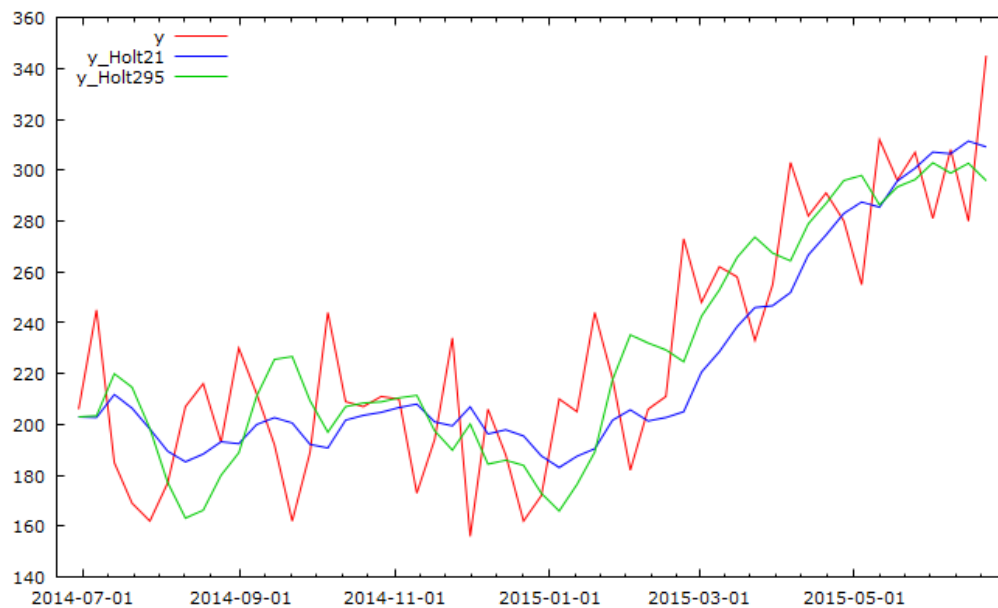
Der Glättungseffekt nimmt ab und die erstellte Zeitreihe reagiert schneller an Niveauänderungen.

Trendschätzung: $b_t = \gamma(L_t - L_{t-1}) + (1 - \gamma)b_{t-1}$

Für γ nahe bei 0

- ✓ die letzte Trendentwicklung erhält eine grössere Gewichtung bei der Trendschätzung, je näher der Glättungsparameter γ an 0 liegt.
- ✓ die länger zurückliegende Trendentwicklung geht stärker in die aktuelle Trendschätzung ein.

Für γ nahe bei 1 reagiert die Zeitreihe schneller an Niveauänderungen ($L_t - L_{t-1}$) → Zeitreihe wird weniger geglättet.



34. Geben Sie für das letzte Menu „number of final intra-sample predictions“ die Zahl 4 ein. Gretl wird für die letzten 4 Beobachtungen 48-52 eine **in-sample** Prognose (innerhalb der Stichprobe) anhand der Information in Periode 47 erstellen.

Argumente auswählen:

Dependent variable (series)	<input type="text" value="y"/>
Level smoothness parameter (scalar)	<input type="text" value="0.2"/>
Slope smoothness parameter (scalar)	<input type="text" value="0.1"/>
Initial values based on:	<input type="text" value="a regression"/>
Number of final intra-sample predictions	<input type="text" value="4"/>

48	307	300,6860
49	281	307,0494
50	308	312,1500
51	280	317,2506
52	345	322,3512

Statistics for the quality of predictions:

	Mean Sq. Err.	Mean. Abs. Err.
HOLT:	789,3656	25,0750

Number of predictions: 4 (49-52)

48	307	301,4232
49	281	307,6721
50	308	312,5447
51	280	317,4173
52	345	322,2899

Statistics for the quality of predictions:

	Mean Sq. Err.	Mean. Abs. Err.
HOLT:	797,5917	25,2724

Number of predictions: 4 (49-52)

35. Fügen Sie in die Zeitreihe **drei** neue Beobachtungen mit den Werten 0 hinzu. Erstellen Sie jetzt eine **out-of-sample** Prognose (ausserhalb der Stichprobe) mit den gleichen Glättungsparameter wie vorher.

The screenshot shows the CAS software interface with several overlapping windows and data tables. A red box highlights the 'Zeige Werte' (Show Values) button in the top left. Another red box highlights the 'y' column in a data table. A third red box highlights the 'y' column in a forecast table. A text box with a plus sign and a checkmark is visible, and a callout box says 'Nach jedem Eintrag Haken betätigen' (Press the checkmark after each entry). The data table shows the following values:

	y
52	345
53	0
54	0
55	0

The forecast table shows the following values:

	Mean
53	206
54	245
55	---

The HOLT statistics are shown as:

HOLT:	109174,1662	330,3928
-------	-------------	----------

The number of predictions is 3 (53-55).