

## Rechenbeispiel mit EViews für Regression mit zwei Regressoren ("Düngerbeispiel" bei L. v. Auer)

### 1. Dateneingabe (Achtung: bei diesem Beispiel handelt es sich um Querschnittsdaten)

Das Rechenbeispiel ist im Lehrbuch von L. v. Auer unter Tabelle 8.2 zu finden und in dem Buch wird hierauf immer wieder Bezug genommen unter den "Numerischen Illustrationen" Nr. 9.1 bis 9.9, aber auch in Kapitel 10. Die Variablen sind wie folgt bezeichnet

PHOS1	$X_{1t} = \ln(p_t) = \text{logarithmierter Phosphat-Input}$
NIT2	$X_{2t} = \ln(n_t) = \text{logarithmierter Stickstoff*-Input}$
GER	$y_t = \text{logarithmierter Gerste-Output}$

\* Nitrogenium = Stickstoff

Als screen shot für die eingegebenen Daten erhält man das folgende Bild:

The screenshot shows the EViews interface. On the left, the 'Workfile: DÜNGERBEISPIE...' window displays a list of variables: c, ger, nit2, phos1, resid, and vonauer. On the right, the 'Group: VONAUER' window shows a data grid with 30 observations for variables PHOS1, NIT2, and GER. The data values are as follows:

obs	PHOS1	NIT2	GER
1	3.091000	3.688900	3.647000
2	3.091000	4.094300	3.892400
3	3.091000	4.499800	4.092200
4	3.091000	4.787500	4.083500
5	3.135500	3.912000	3.816600
6	3.135500	4.382000	3.974600
7	3.135500	4.605200	4.035100
8	3.135500	4.787500	3.930100
9	3.178100	3.688900	3.803800
10	3.178100	4.094300	3.990100
11	3.178100	4.499800	4.100000
12	3.178100	4.787500	4.064100
13	3.218900	3.912000	3.942000
14	3.218900	4.382000	4.070400
15	3.218900	4.605200	4.047800
16	3.218900	4.700500	4.086800
17	3.258100	3.912000	4.011700
18	3.258100	4.248500	3.991400
19	3.258100	4.499800	4.198300
20	3.258100	4.700500	4.122900
21	3.295800	3.688900	3.891600
22	3.295800	4.094300	3.996000
23	3.295800	4.382000	4.065300
24	3.295800	4.605200	4.140100
25	3.332200	3.912000	3.927100
26	3.332200	4.248500	3.942400
27	3.332200	4.605200	4.084100
28	3.332200	4.700500	4.222000
29	3.367300	4.094300	4.081800
30	3.367300	4.605200	4.165000

Workfile-Liste ein Basisbestandteil und eine Art Inhaltsverzeichnis des Workfiles mit allen Objekten" (d.h. Variablen, der Gruppe G, benannten Gleichungen etc). **resid** und **c** (Absolutglied der Regression) sind automatisch vom Programm angelegte Variablen.

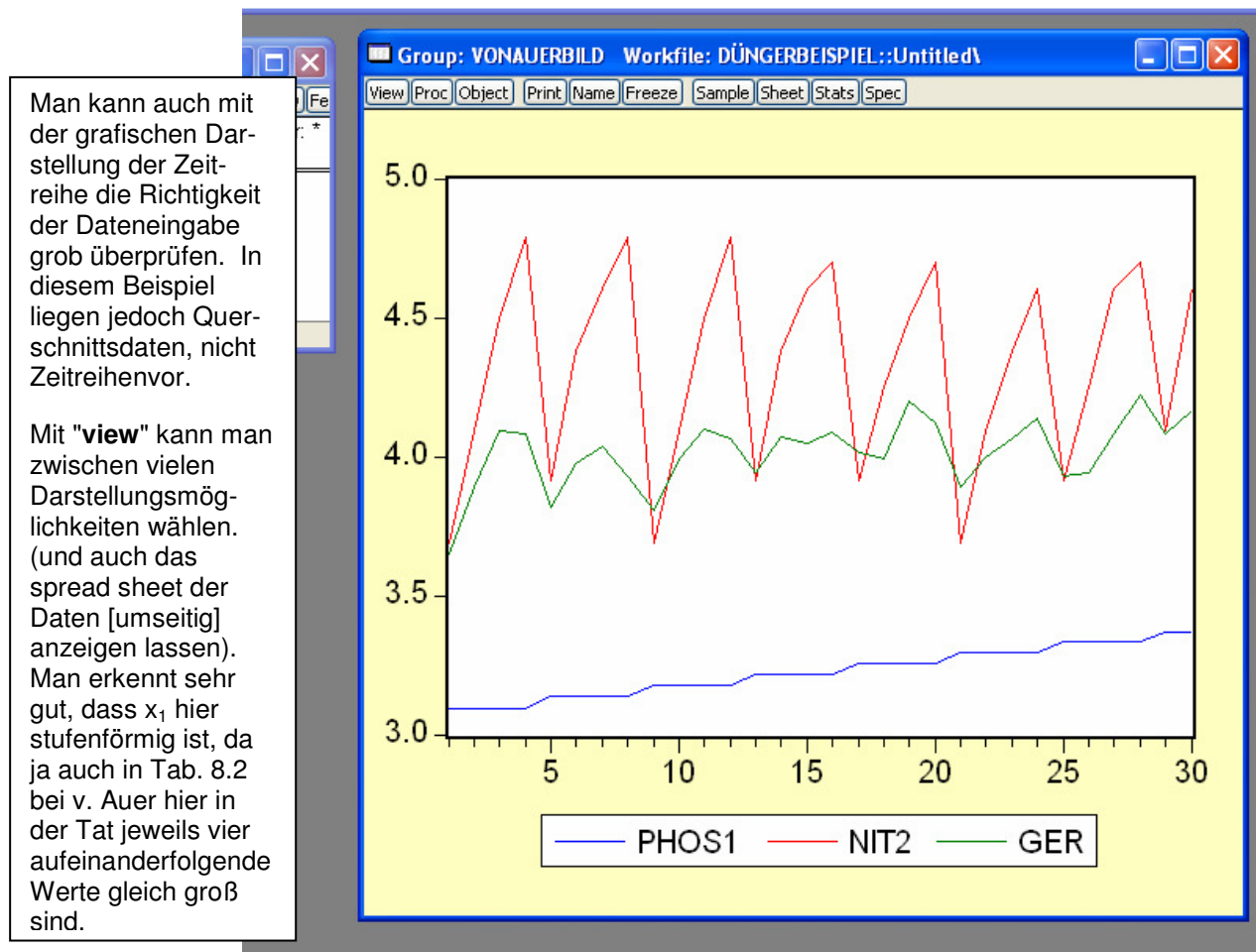
Diese Liste **nicht** während der Arbeit mit dem roten **x** **löschen**.

Bei der (hier manuellen, nicht aus einer Datei übernommenen) Dateneingabe beachten:

Start mit File-New-Workfile(WF) beginnen. Es erscheint ein Fenster "Workfile Create". Dort Art der Daten (Zeitreihen, Querschnittsdaten etc.) eingeben und WF einen Namen geben (hier

"Düngebeispiel"). Dann Object-New Object<sup>1</sup>-Series, die einzelnen Reihen benennen<sup>2</sup> (sonst kann man keine group definieren, denn dort wird eine list of series verlangt [Eingabe der Liste mit blancs]). Bei Zahleneingabe achten auf den "Edit +/-" toggle Schalter: Edit-Modus + heißt Eingabemodus, Edit - ist der protection mode.

## 2. Überprüfung der Eingabe



## 3. Regressionsgleichung

Es gibt mindestens vier Möglichkeiten die Regressionsfunktion zu bestimmen:

1. mit **quick** (obere Befehlsleiste) - generate series by equation
2. mit **genr** oder
3. mit **proc** (untere Befehlsleiste) dann "make equation" oder über
4. **object** (unten) – new object – equation (beste Möglichkeit)

Die Gleichung selbst wird wie folgt eingegeben (mit blancs ohne Komma und = Zeichen)

```
ger c phos1 nit2
```

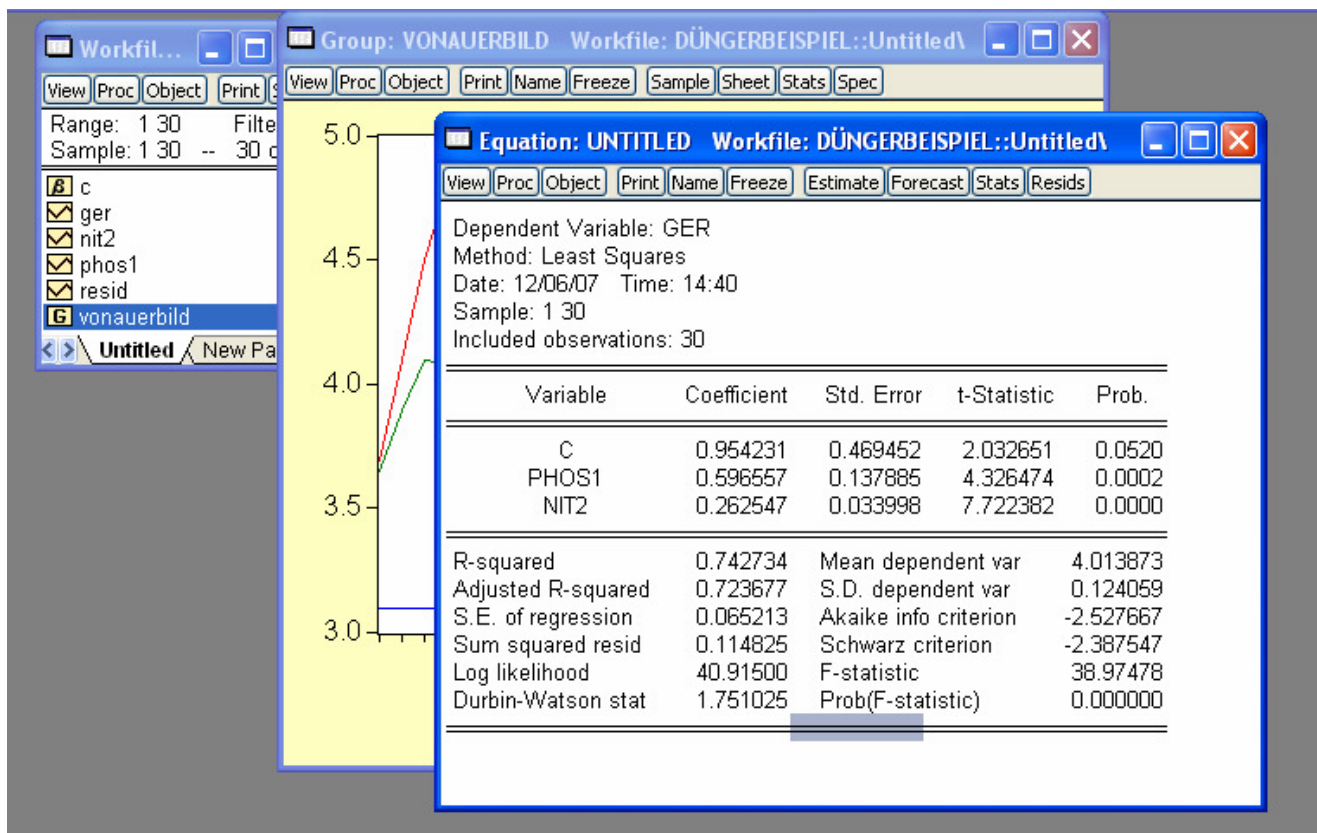
Man erkennt jetzt auf dem screen shot umseitig die in der "Numerischen Illustration 9.1" des Lehrbuchs von v. Auer mitgeteilten Werte wieder (mit geringen Abweichungen):

<sup>1</sup> Ein object ist in EView eine Zeitreihe, eine Gruppe (group, also mehrere Zeitreihen), eine Gleichung, Graphik, Tabelle oder auch ein Text etc.

<sup>2</sup> Default ist series01, series02,...Name nach Eingabe wählen mit "name".

$\hat{\alpha} = 0,95423$   $\hat{\beta}_1 = 0,59652$  (hier 0,596557 statt 0,59652) und  $\hat{\beta}_2 = 0,26255$ .

Wie man hier sieht, haben der Workfile und die Graphik einen Namen, die Gleichung aber nicht. Die drei Variablen (ger, phos1 und nit2) gehören zur group "vonauer".



Man erkennt auch im Computer-Ausdruck leicht die in der "Numerischen Illustration 9.2" angegebenen Werte  $R^2 = 74,3\%$  (0,742734) sowie  $S_{\hat{u}\hat{u}} = 0,11481$ .  $S_{\hat{u}\hat{u}}$  ist die **sum of squared residuals** (0,114825). Mit dieser Größe kann man die Varianzzerlegung (F-Test) und die **Standardabweichung der Schätzung** (S.E. of regression), als Symbol  $\hat{\sigma}$ , über die geschätzte Varianz der Störgröße  $\hat{\sigma}^2$  wie folgt errechnen  $\hat{\sigma}^2 = \frac{S_{\hat{u}\hat{u}}}{T - K - 1}$ , wobei hier  $T - K - 1 = 30 - 2 - 1 = 27$  ist. Wie man sieht

ist  $\hat{\sigma}^2 = 0,114825/27 = 0,00425278$  und damit ist  $\hat{\sigma} = 0,0652133$  (der [standard error] S.E.of regression).

Der "equation-output"<sup>3</sup> zeigt auch die (geschätzten) **Standardabweichungen der Regressionskoeffizienten**, aus denen sich dann die (geschätzten) Varianzen der Koeffizienten (vgl. "Numerische Illustration 9.4") errechnen lassen

$\hat{\sigma}_{\hat{\alpha}}^2 = (0,469452)^2 = 0,220385$  (bei v. Auer 0,22037),  $\hat{\sigma}_{\hat{\beta}_1}^2 = (0,137885)^2 = 0,0190123$  (bei v.A.: 0,01901) und  $\hat{\sigma}_{\hat{\beta}_2}^2 = (0,033998)^2 = 0,00115586$  (bei von Auer: 0,00116).

Wie man an den **t-Werten** (und **prob values**) sieht ist  $\alpha$  nicht auf dem 5% Niveau gesichert, die anderen Koeffizienten sind dagegen hochsignifikant.

Der von EViews errechnete **F-Wert** für das Testen der Hypothese  $H_0: \beta_1 = \beta_2 = 0$  in Höhe von 38,97478 findet sich auch als  $\approx 38,98$  bei v. Auer (siehe Numerische Illustration 10.9). Auch hier zeigt der prob-value ("Prob(F statistic)") in Höhe von 0,000000, dass  $H_0$  (Annahme der Irrelevanz beider Regressoren) mit weniger als 1% (also 0,001) Irrtumswahrscheinlichkeit (Signifikanzniveau) verworfen (rejected) werden kann [denn  $0,001 > \text{Prob}(F \text{ statistic}) = 0,000000$ ] abgelehnt. Die

<sup>3</sup> Man erkennt auch das Datum, an dem ich die Berechnung durchgeführt habe (6. Dez. 2007 um 14:40 Uhr).

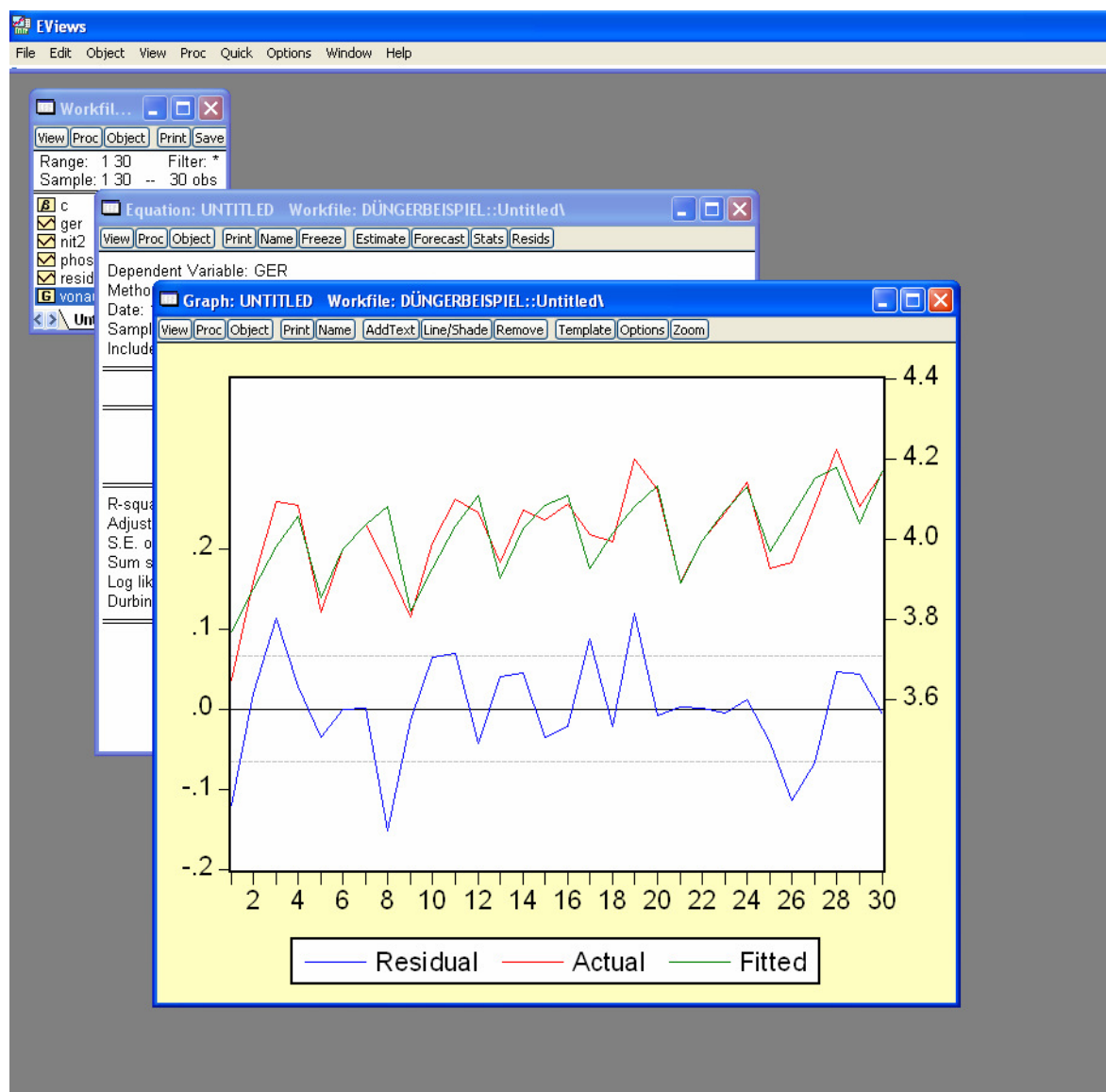
Einflussgrößen PHOS ( $x_1$ ) und NIT ( $x_2$ ) sind also *einzel*n (t-Test) *und in ihrer Gesamtheit* (F-Test) signifikant (nicht irrelevant).

#### 4. Ergänzungen

Die Funktionen Estimate<sup>4</sup>, Forecast und Resids bieten weitere Möglichkeiten, von denen im Folgenden nur auf die Graphik zur Beurteilung der Anpassung und der Verteilung der (geschätzten) Störgröße (also der Residuen) eingegangen werden soll.

Weitere Informationen über die Variablen (z.B. Mittelwerte, Varianzen) und auch über die Korrelation der Regressoren untereinander erhält man mit der Taste **View**

**Resids** liefert die folgende Graphik<sup>5</sup>



In diesem Beispiel ist die Graphik der Residuen und insbesondere die Funktion **Forecast** nicht sonderlich instruktiv, weil wir es hier mit Querschnittsdaten (Versuchsfelder Nr. 1 bis 30) und nicht mit Zeitreihen zu tun haben.

<sup>4</sup> Hier kann man z.B. Korrekturen wegen Heteroskedastizität (Newey West Schätzung usw.) vornehmen.

<sup>5</sup> Die Graphik ist nicht benannt (Graph untitled), wollte man mehrere Graphiken machen und sie auch weiter bearbeiten (siehe die hier relevanten Befehle, z.B. "add text") und nach Beendigung der Sitzung erhalten lassen wäre es besser, ihr einen Namen zu geben.