



Make Big Money

各國匯率預測
資料探勘技術應用

專題 F 組

指導教授 趙景明

97156201 竇瑋蓉

97156241 陳思穎

97156255 周鈺祥

報告大綱

- ⊕ 研究動機
- ⊕ 研究目的
- ⊕ 軟體工具
- ⊕ 研究流程

- ⊕ 成果展現
- ⊕ 研究結論



研究動機



研究目的

- ⊕ 藉由資料探勘技術建立預測模型
- ⊕ 預測出各國匯率值
- ⊕ 從模型中了解影響各國匯率的因子
- ⊕ 幫助投資人提高投資報酬率



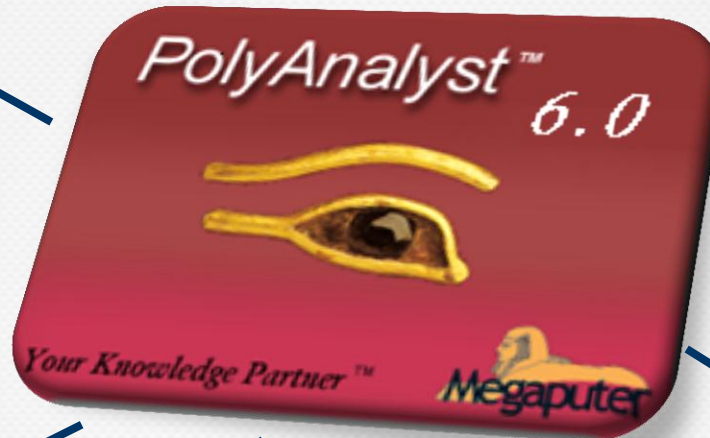
軟體工具

支援完整Data mining

具有Client/ Server架構

支援微軟資料倉儲聯盟
通訊介面OLE DB

完整前後處理功能



豐富的資料處理工具

能透過SQL 查詢語法
評分資料倉儲內資料

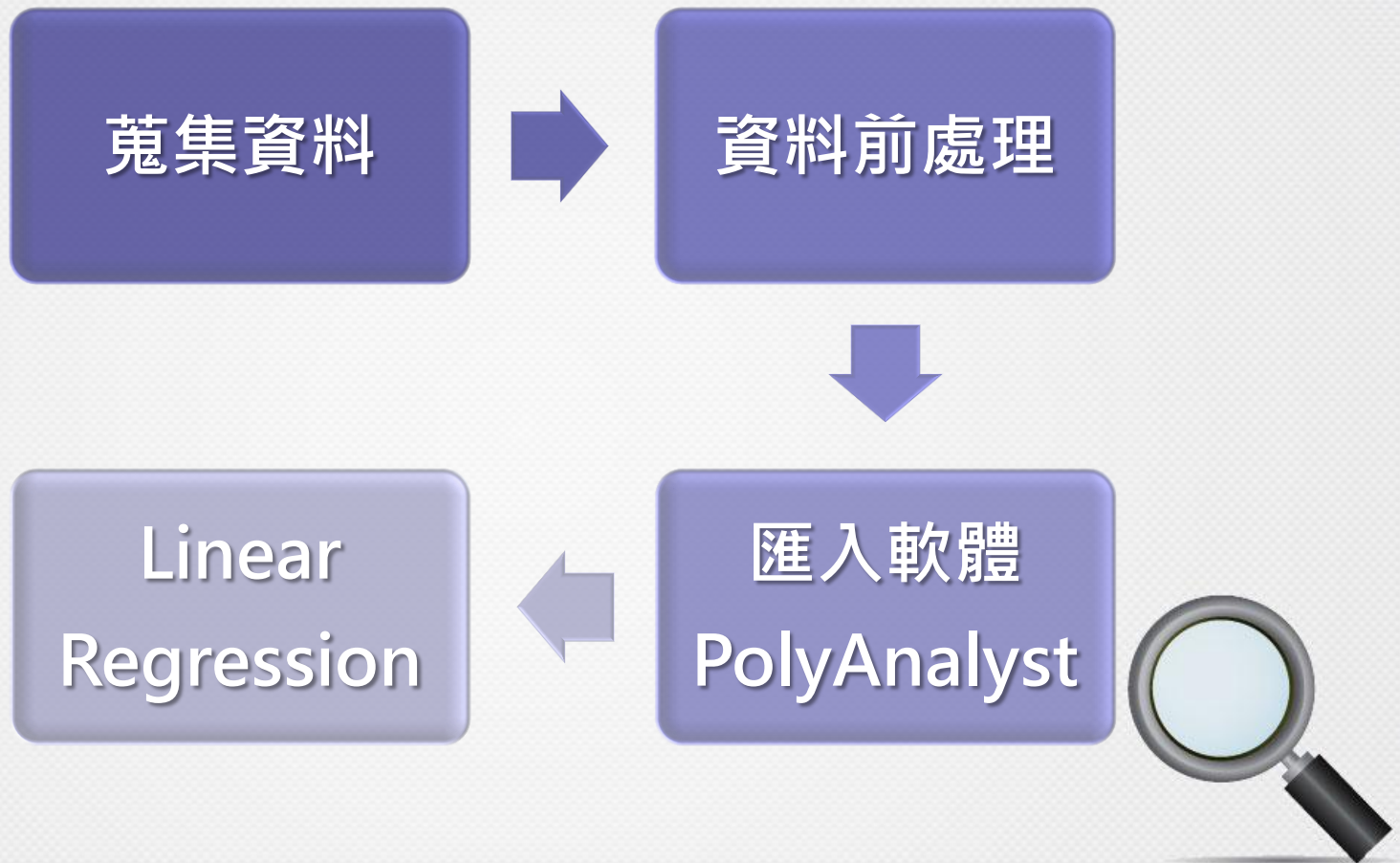
和善的視覺化介面

具Multithreaded 演算法
做為資料平行處理

解決多維度相依問題
具有24種機器學習演算法



研究流程



匯率

蒐集資料

加拿大元
墨西哥比索

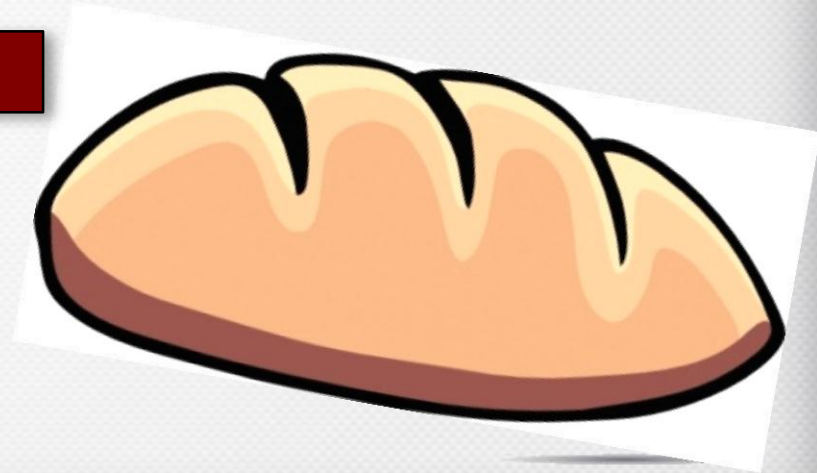
挪威克朗
歐元
英鎊
法郎

辛巴威

新台幣
人民幣
港幣
日圓
韓元

新加坡
紐西蘭

辛巴威



蒐集資料

東京 (日經225)
韓國Electric Power
Corporatio

香港 (恆生指數)
上海SINOPEC Shangai
Petrochemical

各國大盤

蘇黎士
法蘭克福 (商銀指數)
倫敦 (金融時報指數)

NASDAQ
SOX INDU

⊕ 資料擷取範圍：2002.04-2011.12 單位：每日



蒐集資料

Pan American Silver Corp

Silver Standard Resources

National Oilwell Varco, Inc. Co

ATP Oil & Gas Corporation: 美國休斯頓石油天然氣生產商

Corn Products International: 美國伊利諾斯州玉米提煉廠

Green Mountain Coffee Roasters : 美國特種咖啡生產商綠山咖啡

Peet's Coffee & Tea, Inc. (PEET): 以加州為主的大型連鎖咖啡店

The Cheesecake Factory Incorpor: 加州連鎖餐廳提供乳酪蛋糕食品

Southwest Airlines Company Comm (LUV)

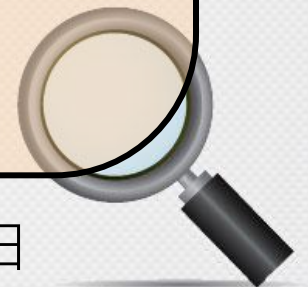
Overseas Shipholding Group, Inc (OSG)

Panera Bread Company (PNRA)

American Woodmark Corporation (AMWD)

Rocky Mountain Chocolate Factor (RMCF): 洛磯山脈巧克力工廠

⊕ 資料擷取範圍：2002.04-2011.12 單位：每日



資料前處理

Date	美元對 英鎊匯 率	美元對 墨西哥 比索	美元 兌韓 元匯 率	美元 兌日 元匯 率	美元 兌瑞 士法 郎匯 率	美元 兌紐 西蘭 元匯 率	美元 兌新 台幣 匯率	美元兌人 民幣匯率	美元 兌新 加坡 元匯 率	美元 兌加 拿大 元匯 率	美元 兌新 加坡 元匯 率	美元 兌歐 元匯 率	美元 兌港 幣匯 率	NASDAQ Q1000 Index	SONIC Close	INDUCE Fee	America S	Silver Standard	National Oilwell	ATF Oil & Gas	Corn Product
2002年4月2日	0.6965	9.002	1329	133.4	1.666	2.262	35.01	8.277	8.755	1.593	1.847	1.139	7.8	1410.41	597.29	10.110.71	5.46	3.1	25.85	4.19	31
2002年4月3日	0.6969	9.04	1329	132.65	1.661	2.276	35	8.277	8.692	1.589	1.842	1.136	7.8	1394.21	574.21	10.108.23	5.29	2.96	24.89	3.9	31
2002年4月4日	0.6969	9.048	1330	132.3	1.665	2.274	35	8.277	8.715	1.593	1.842	1.139	7.8	1397.91	583.91	10.108.71	5.32	2.98	24.91	3.9	31
2002年4月5日	0.6988	9.021	1328	131.27	1.674	2.312	34.98	8.277	8.718	1.594	1.836	1.143	7.8	1395.21	576.21	10.108.23	5.29	2.96	24.89	3.9	31
2002年4月8日	0.6971	9.042	1330	131.03	1.668	2.302	34.98	8.277	8.688	1.596	1.837	1.137	7.8	1396.21	581.21	10.108.71	5.31	2.97	24.90	3.9	31
2002年4月10日	0.6961	9.05	1329	130.75	1.67	2.288	34.98	8.277	8.706	1.593	1.835	1.137	7.8	1395.21	576.21	10.108.23	5.29	2.96	24.89	3.9	31
2002年4月11日	0.6953	9.074	1330	131.55	1.66	2.259	34.99	8.277	8.636	1.588	1.838	1.133	7.8	1396.21	581.21	10.108.71	5.31	2.97	24.90	3.9	31
2002年4月12日	0.6958	9.078	1332	132.09	1.668	2.261	35	8.277	8.654	1.588	1.838	1.137	7.8	1351.85	560.67	10.108.23	5.2	2.84	23.22	3.5	31
2002年4月15日	0.6958	9.182	1328	131.95	1.667	2.271	34.99	8.277	8.667	1.589	1.837	1.136	7.8	1353.21	563.21	10.108.23	5.21	2.85	23.23	3.5	31
2002年4月16日	0.6946	9.162	1326	131.19	1.664	2.274	34.98	8.277	8.658	1.58	1.834	1.132	7.8	1352.21	562.21	10.108.23	5.2	2.84	23.22	3.5	31
2002年4月17日	0.6926	9.205	1319	130.7	1.655	2.252	34.96	8.277	8.607	1.574	1.829	1.126	7.8	1349.21	559.21	10.108.23	5.19	2.83	23.21	3.5	31
2002年4月18日	0.6907	9.242	1313	129.84	1.648	2.245	34.94	8.277	8.585	1.575	1.824	1.124	7.8	1347.21	557.21	10.108.23	5.18	2.82	23.20	3.5	31
2002年4月19日	0.691	9.254	1312	130.5	1.651	2.239	34.9	8.278	8.555	1.572	1.822	1.124	7.8	1346.21	556.21	10.108.23	5.18	2.82	23.20	3.5	31
2002年4月22日	0.6902	9.254	1309	129.9	1.654	2.241	34.85	8.277	8.573	1.573	1.814	1.126	7.8	1349.83	573	10.126.43	5.75	2.83	23.91	3.75	31
2002年4月23日	0.6903	9.27	1306	129.89	1.649	2.239	34.81	8.277	8.54	1.575	1.818	1.124	7.8	1348.83	571	10.126.43	5.74	2.82	23.9	3.75	31
2002年4月24日	0.6906	9.286	1307	129.76	1.644	2.232	34.81	8.277	8.543	1.568	1.816	1.124	7.8	1348.83	571	10.126.43	5.74	2.82	23.9	3.75	31
2002年4月25日	0.6878	9.36	1298	128.4	1.631	2.219	34.78	8.277	8.458	1.56	1.81	1.124	7.8	1347.83	569	10.126.43	5.73	2.81	23.89	3.75	31
2002年4月26日	0.6875	9.319	1298	128.18	1.629	2.224	34.77	8.277	8.441	1.563	1.808	1.124	7.8	1346.83	567	10.126.43	5.72	2.81	23.88	3.75	31
2002年4月30日	0.6866	9.375	1294	128.45	1.622	2.236	34.72	8.277	8.405	1.556	1.808	1.124	7.8	1345.83	565	10.126.43	5.71	2.80	23.87	3.75	31
2002年5月2日	0.6828	9.435	1288	128.02	1.61	2.245	34.65	8.277	8.382	1.562	1.808	1.124	7.8	1344.83	563	10.126.43	5.70	2.80	23.86	3.75	31
2002年5月7日	0.6814	9.476	1280	127.8	1.593	2.227	34.57	8.277	8.336	1.57	1.808	1.124	7.8	1343.83	561	10.126.43	5.69	2.79	23.85	3.75	31
2002年5月8日	0.6881	9.413	1285	128.64	1.61	2.233	34.59	8.278	8.378	1.572	1.808	1.124	7.8	1346.83	565	10.126.43	5.71	2.80	23.87	3.75	31
2002年5月9日	0.6886	9.444	1282	128.36	1.6	2.216	34.61	8.277	8.306	1.569	1.81	1.1	7.799	1340.41	503.48	10.037.42	5.6	2.78	23.84	3.75	31
2002年5月10日	0.6847	9.512	1280	127.89	1.595	2.21	34.57	8.277	8.302	1.568	1.805	1.097	7.799	1188.78	480.96	9.938.82	6.37	3.21	26.53	4.22	31
2002年5月13日	0.6855	9.488	1277	127.69	1.598	2.197	34.55	8.277	8.267	1.558	1.803	1.096	7.799	1186.78	479.96	9.938.82	6.37	3.21	26.53	4.22	31
2002年5月14日	0.6909	9.525	1279	128.66	1.616	2.196	34.58	8.277	8.365	1.559	1.808	1.108	7.799	1187.78	480.96	9.938.82	6.37	3.21	26.53	4.22	31
2002年5月15日	0.6869	9.455	1278	127.92	1.605	2.186	34.57	8.277	8.323	1.556	1.806	1.102	7.799	1186.78	479.96	9.938.82	6.37	3.21	26.53	4.22	31
2002年5月16日	0.6865	9.49	1270	128.19	1.599	2.168	34.54	8.277	8.294	1.555	1.806	1.098	7.799	1185.78	478.96	9.938.82	6.36	3.20	26.52	4.21	31
2002年5月17日	0.6848	9.478	1262	126.03	1.581	2.158	34.46	8.277	8.202	1.542	1.797	1.086	7.799	1184.78	477.96	9.938.82	6.35	3.19	26.51	4.20	31
2002年5月20日	0.6855	9.494	1254	125.64	1.578	2.148	34.41	8.277	8.186	1.542	1.792	1.085	7.799	1183.78	476.96	9.938.82	6.34	3.18	26.50	4.19	31
2002年5月21日	0.6864	9.51	1250	123.83	1.58	2.141	34.37	8.277	8.194	1.539	1.789	1.088	7.799	1182.78	475.96	9.938.82	6.33	3.17	26.49	4.18	31
2002年5月22日	0.6854	9.535	1242	124.16	1.569	2.121	34.34	8.277	8.072	1.536	1.792	1.079	7.8	1181.78	474.96	9.938.82	6.32	3.16	26.48	4.17	31
2002年5月23日	0.6875	9.527	1241	124.84	1.576	2.1	34.33	8.277	8.073	1.534	1.8	1.083	7.8	1180.78	473.96	9.938.82	6.31	3.15	26.47	4.16	31
2002年5月24日	0.686	9.562	1238	124.71	1.583	2.124	34.36	8.277	8.072	1.535	1.799	1.085	7.8	1179.78	472.96	9.938.82	6.30	3.14	26.46	4.15	31
2002年5月25日	0.6841	9.56	1232	124.38	1.574	2.1	34.28	8.277	7.952	1.53	1.794	1.077	7.8	1178.78	471.96	9.938.82	6.29	3.13	26.45	4.14	31
2002年5月29日	0.6849	9.61	1229	124.35	1.576	2.101	34.21	8.277	7.945	1.534	1.793	1.074	7.8	1177.78	470.96	9.938.82	6.28	3.12	26.44	4.13	31
2002年5月30日	0.6824	9.713	1222	123.08	1.562	2.095	34.06	8.277	7.929	1.534	1.786	1.067	7.8	1176.78	469.96	9.938.82	6.27	3.11	26.43	4.12	31
2002年5月31日	0.6838	9.652	1219	124.13	1.569	2.088	34.05	8.276	8.002	1.528	1.788	1.071	7.8	1175.78	468.96	9.938.82	6.26	3.10	26.42	4.11	31

■ 共40種資料來源

■ 每筆資料來源有3000多筆數據



■ 挑完有將近10萬筆數據



匯入軟體PolyAnalyst

Microsoft Excel Properties

Select Excel file (*.xls;*.xlsx) Column specifications Sample Column folders General

@user@\2002-2006.xlsx

Append data

What to import

Existing worksheet Manually entered SQL expression

Source object/SQL statement

Worksheet: Sheet1\$

2002-2006.xlsx

Linear Regression

Preview

Refresh

Date	1.5 美元對英鎊匯率	1.5 美元兌墨西哥比索匯率	1.5 美元兌韓國韓元匯率	1.5 美元兌日元匯率	1.5 美元兌瑞士法郎匯率	1.5 美元兌紐西蘭匯率	1.5 美元兌新台幣匯率	1.5 美元兌人民幣匯率	1.5 美元
4/2/2002	0.70	9.00	1,329.00	133.40	1.67	2.26	35.01	8.28	
4/3/2002	0.70	9.04	1,329.00	132.65	1.66	2.28	35.00	8.28	
4/4/2002	0.70	9.05	1,330.00	132.30	1.67	2.27	35.00	8.28	
4/8/2002	0.70	9.02	1,328.00	131.27	1.67	2.31	34.98	8.28	
4/9/2002	0.70	9.04	1,330.00	131.03	1.67	2.30	34.98	8.28	
4/10/2002	0.70	9.05	1,329.00	130.75	1.67	2.29	34.98	8.28	
4/11/2002	0.70	9.07	1,330.00	131.55	1.66	2.26	34.99	8.28	
4/12/2002	0.70	9.08	1,332.00	132.09	1.67	2.26	35.00	8.28	
4/15/2002	0.70	9.18	1,328.00	131.95	1.67	2.27	34.99	8.28	
4/16/2002	0.69	9.16	1,326.00	131.19	1.66	2.27	34.98	8.28	
4/17/2002	0.69	9.21	1,319.00	130.70	1.66	2.25	34.96	8.28	
4/18/2002	0.69	9.24	1,313.00	129.84	1.65	2.25	34.94	8.28	
4/19/2002	0.69	9.22	1,312.00	130.50	1.65	2.24	34.90	8.28	
4/22/2002	0.69	9.25	1,309.00	129.90	1.65	2.24	34.85	8.28	
4/23/2002	0.69	9.27	1,306.00	129.89	1.65	2.24	34.81	8.28	
4/24/2002	0.69	9.29	1,307.00	129.76	1.64	2.23	34.81	8.28	
4/25/2002	0.69	9.36	1,298.00	128.40	1.63	2.22	34.78	8.28	
4/26/2002	0.69	9.32	1,298.00	128.18	1.63	2.22	34.77	8.28	
4/30/2002	0.69	9.38	1,294.00	128.45	1.62	2.24	34.72	8.28	
5/2/2002	0.68	9.44	1,288.00	128.02	1.61	2.25	34.65	8.28	
5/7/2002	0.68	9.48	1,280.00	127.80	1.59	2.23	34.57	8.28	
5/8/2002	0.69	9.41	1,285.00	128.64	1.61	2.23	34.59	8.28	
5/9/2002	0.69	9.44	1,282.00	128.36	1.60	2.22	34.61	8.28	
5/10/2002	0.68	9.51	1,280.00	127.89	1.60	2.21	34.57	8.28	
5/13/2002	0.69	9.49	1,277.00	127.69	1.60	2.20	34.55	8.28	
5/14/2002	0.69	9.53	1,279.00	128.66	1.62	2.20	34.58	8.28	
5/15/2002	0.69	9.46	1,278.00	127.92	1.61	2.19	34.57	8.28	
5/16/2002	0.69	9.49	1,270.00	128.19	1.60	2.17	34.54	8.28	
5/17/2002	0.68	9.48	1,262.00	126.03	1.58	2.16	34.46	8.28	
5/20/2002	0.69	9.49	1,254.00	125.64	1.58	2.15	34.41	8.28	

Linear Regression

- Data Analysis
 - Basket Analysis
 - Transactional Basket Analysis
 - Case-Based
 - Neural Network
 - Bayes Network
 - Data Audit
 - CHAID
 - Discriminant Basket Analysis
 - Clustering
 - Decision Tree
 - Find Laws
 - Instance-Based
 - Linear Regression**
 - MediCop
 - MediCop Sequential Analysis
 - Logistic Regression
 - SVM
 - Factor analysis
 - Random Forest
 - Temporal Associations
 - Link Analysis

Linear Regression Properties

Select columns | Multiple targets | Settings | General

Available columns:

Name	Folder
Date	
1.5 Corn Products Internatio...	
1.5 HSBC HLDG (HSBA#L)-L...	
1.5 Corinthian Colleges, Inc#...	
1.5 Green Mountain Coffee ...	
1.5 pb	

Dependent columns:

1.5 美元兌歐元匯率

Independent columns:

Name	Folder
1.5 美元兌日元匯率	
1.5 美元兌瑞士法郎匯率	
1.5 美元兌紐西蘭匯率	
1.5 美元兌新台幣匯率	
1.5 美元兌人民幣匯率	
1.5 美元兌挪威克朗匯率	
1.5 美元兌加拿大元匯率	
1.5 美元兌新加坡幣匯率	
1.5 美元兌港幣匯率	
1.5 NASDAQ100Close	
1.5 SOXClose	
1.5 Pan American Silver Corp...	
1.5 National Oilwell Varco, In...	
1.5 ATP Oil & Gas Corporatio...	
1.5 Peet's Coffee & Tea, Inc...	

Execute OK Cancel

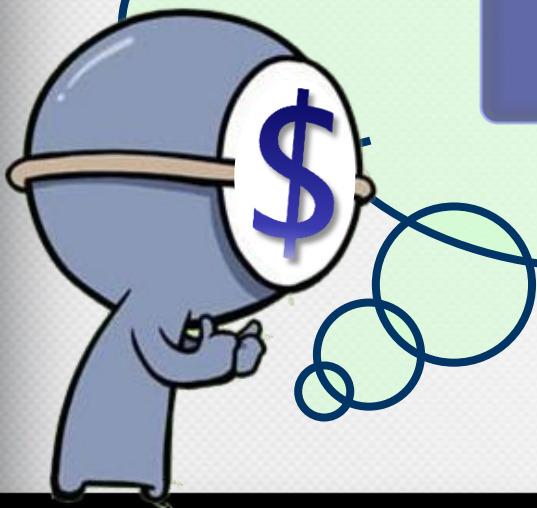
研究方法一

將資料分成訓練資料和測試資料

02-06年為訓練資料 07-11年為測試資料

拿訓練資料得到的公式驗證測試資料

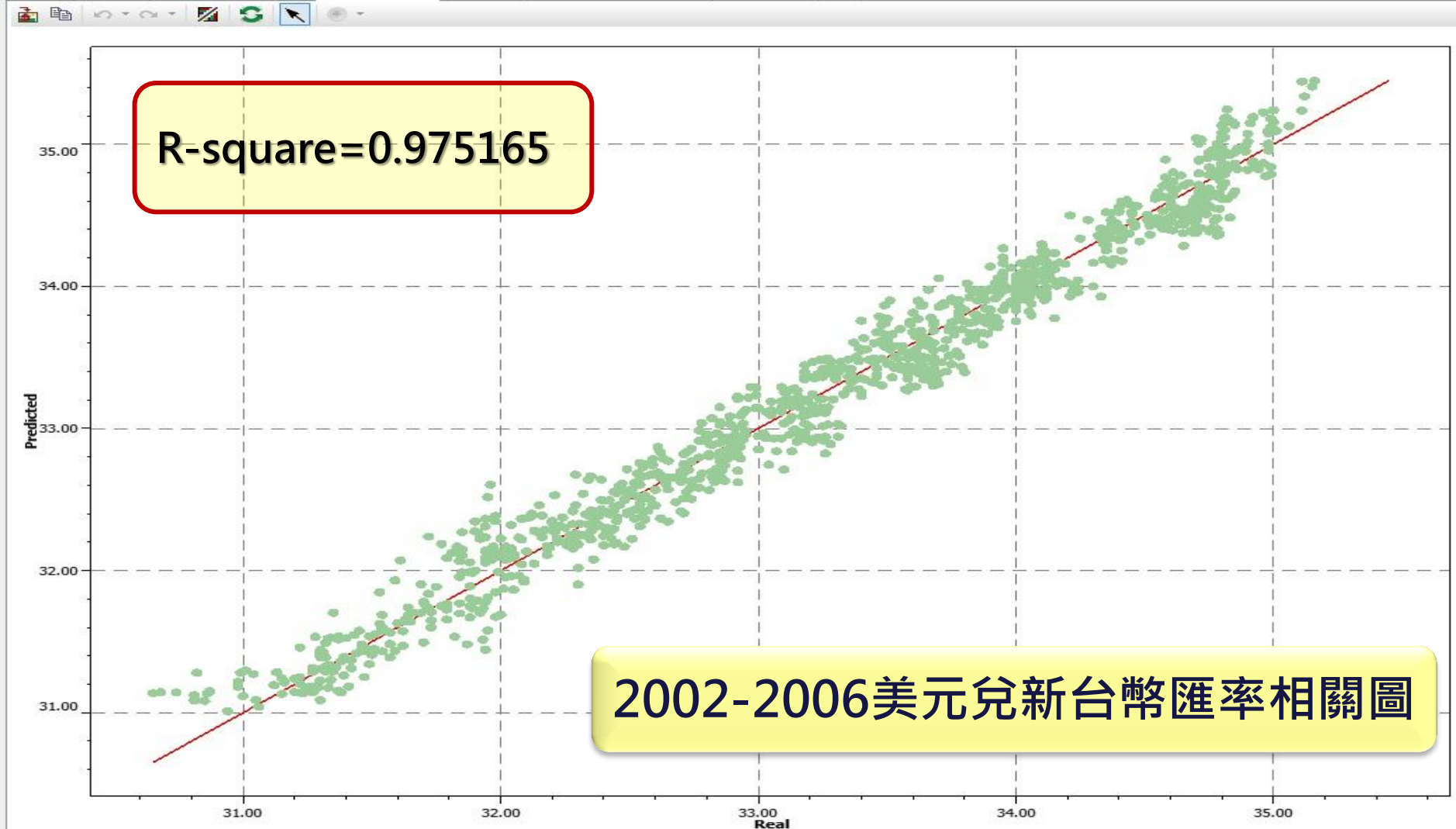
成功賺大錢!!



研究方法一：美元兌新台幣

Target: 美元兌新台幣匯率

Settings Model Prediction Accuracy Pred. vs. Real Predicted, real vs. record number Terms description



2002-2006美元兌新台幣匯率相關圖

2007-2011美元兌新台幣匯率差距圖

- Legend
- 美元兌新台幣匯率
 - 美元兌新台幣預測

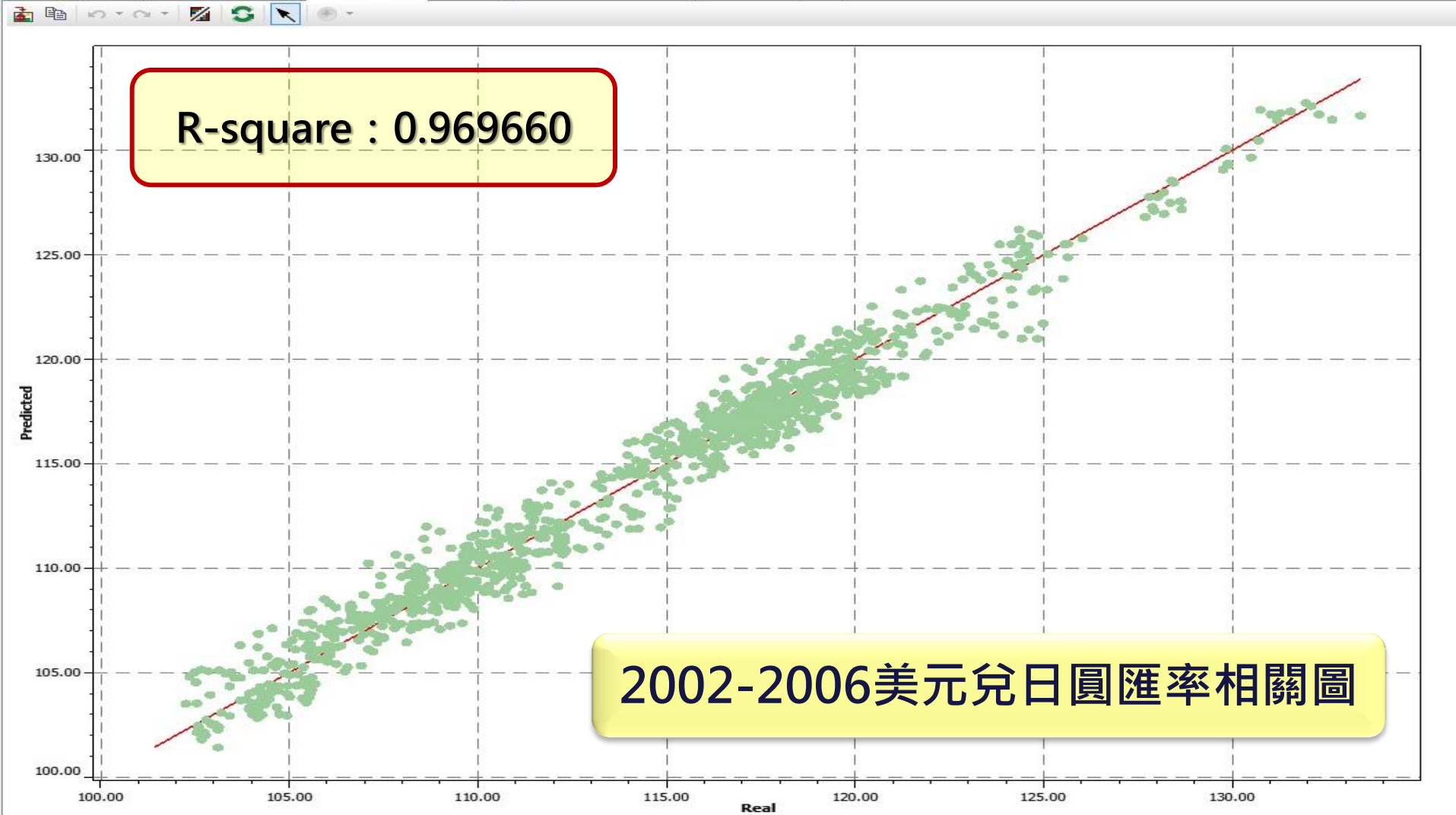


預測差很大!!

研究方法一：美元兌日圓

Target: 美元兌日圓匯率

Settings Model Prediction Accuracy Pred. vs. Real Predicted, real vs. record number Terms description





Legend

- 美元兌日元匯率
- 美元兌日元預測

2007-2011美元兌日圓匯率差距圖



預測差很大!!

誤差：08年金融海嘯

- ⊕ 問題：發現預測資料和實際資料自2008年開始有明顯的變動差距
- ⊕ 2008年發生金融海嘯，是繼1929年經濟大蕭條以來最慘烈的衰退，美國接連發生的次級房貸、雷曼兄弟倒閉等的金融風暴，不僅僅對美國，對於世界經濟體系也造成影響。
- ⊕ “The dollar is our currency, but your problem”
- ⊕ 美國印鈔問題

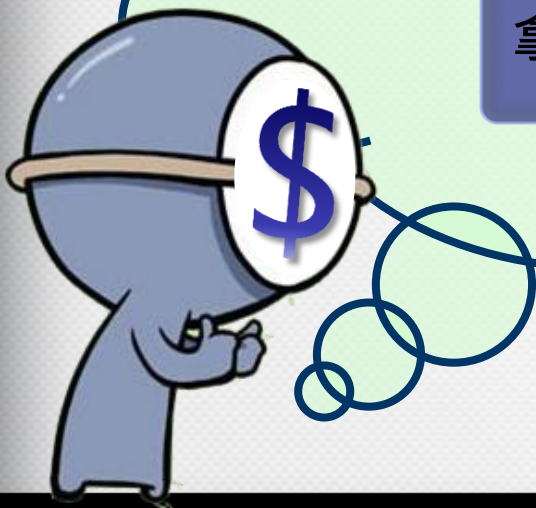


研究方法二

將資料範圍縮小

09-10年為訓練資料 11年為測試資料

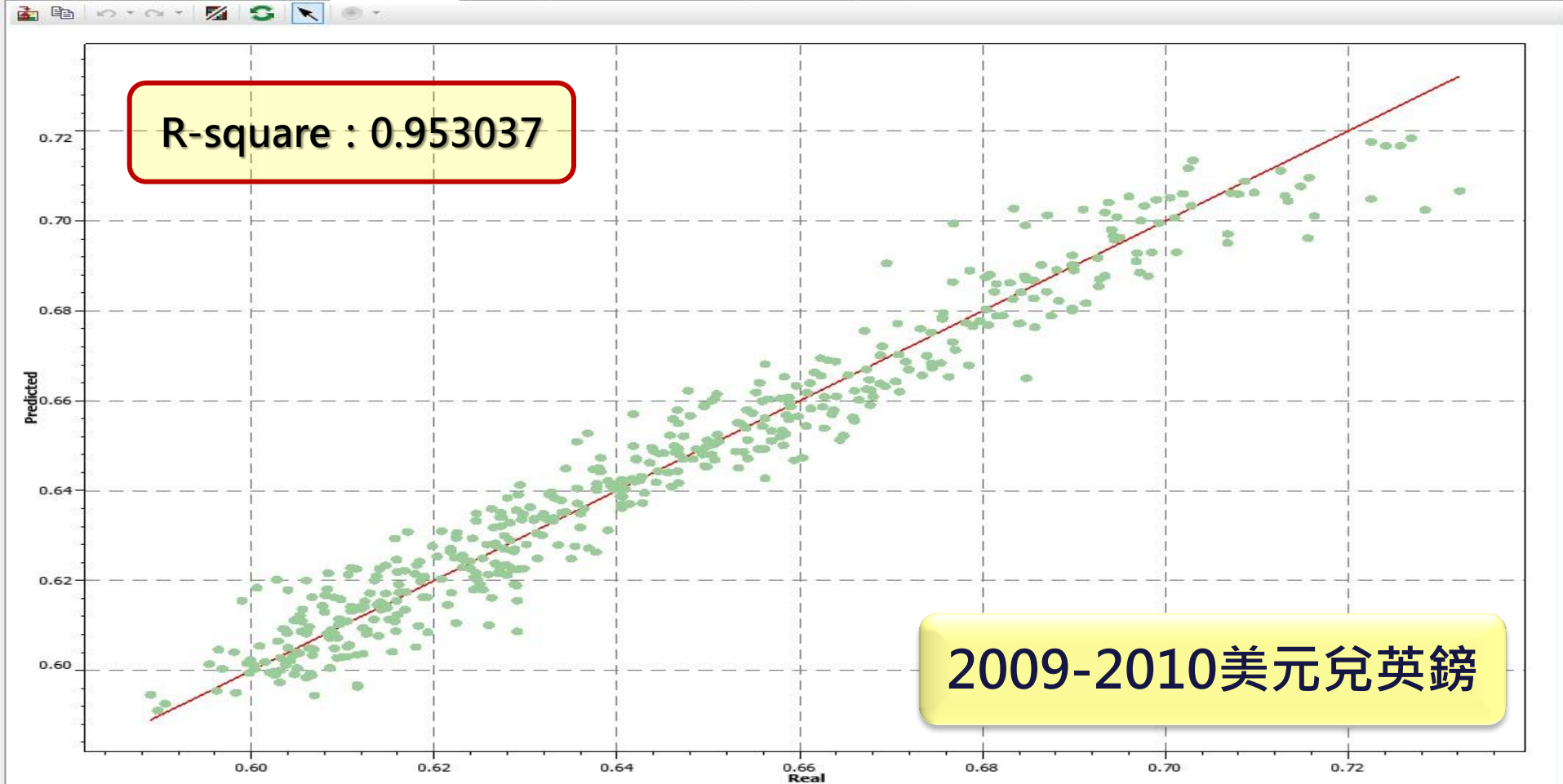
拿訓練資料得到的公式驗證測試資料



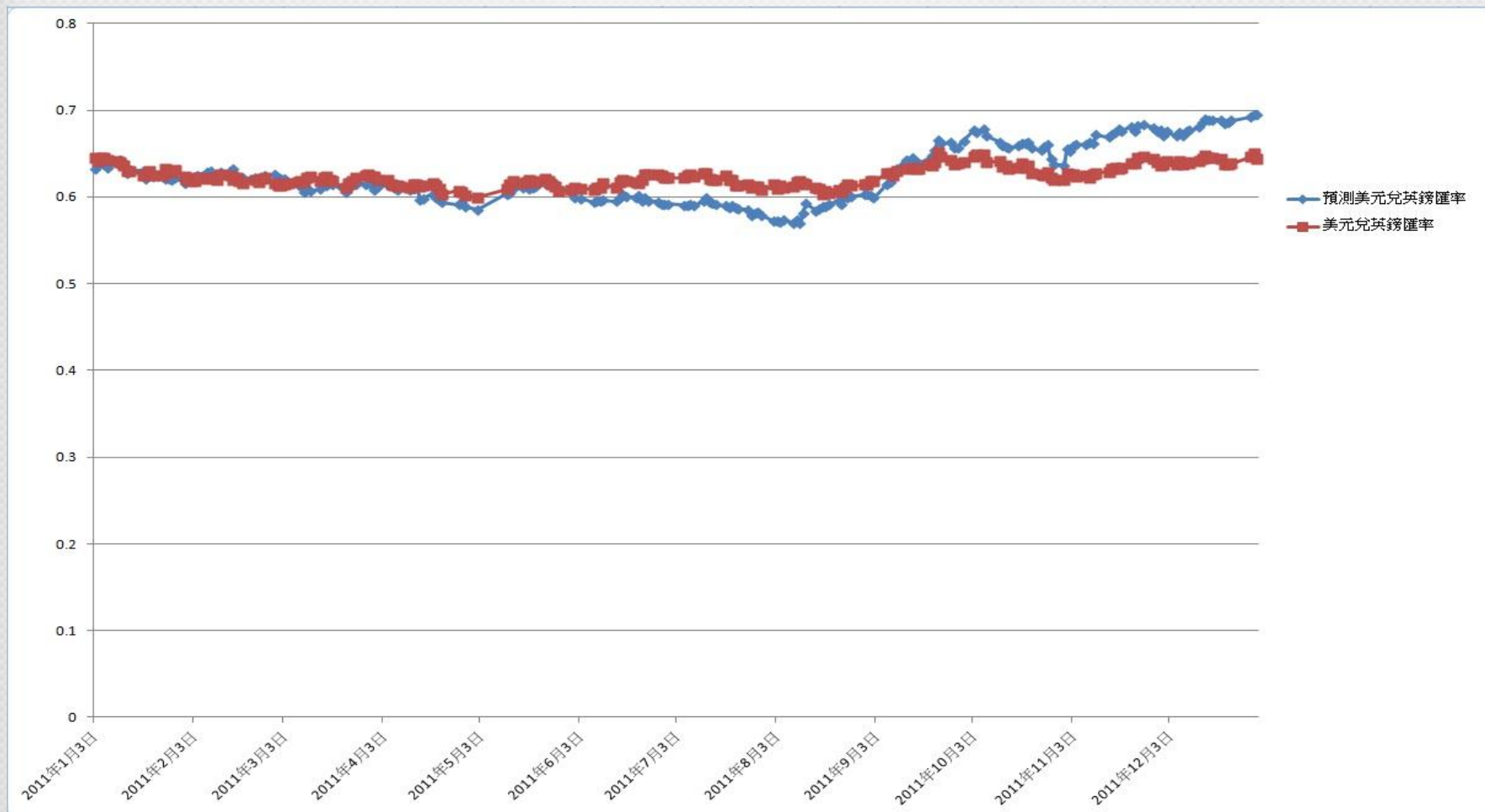
研究方法二：美元兌英鎊

Target: 美元兌英鎊匯率

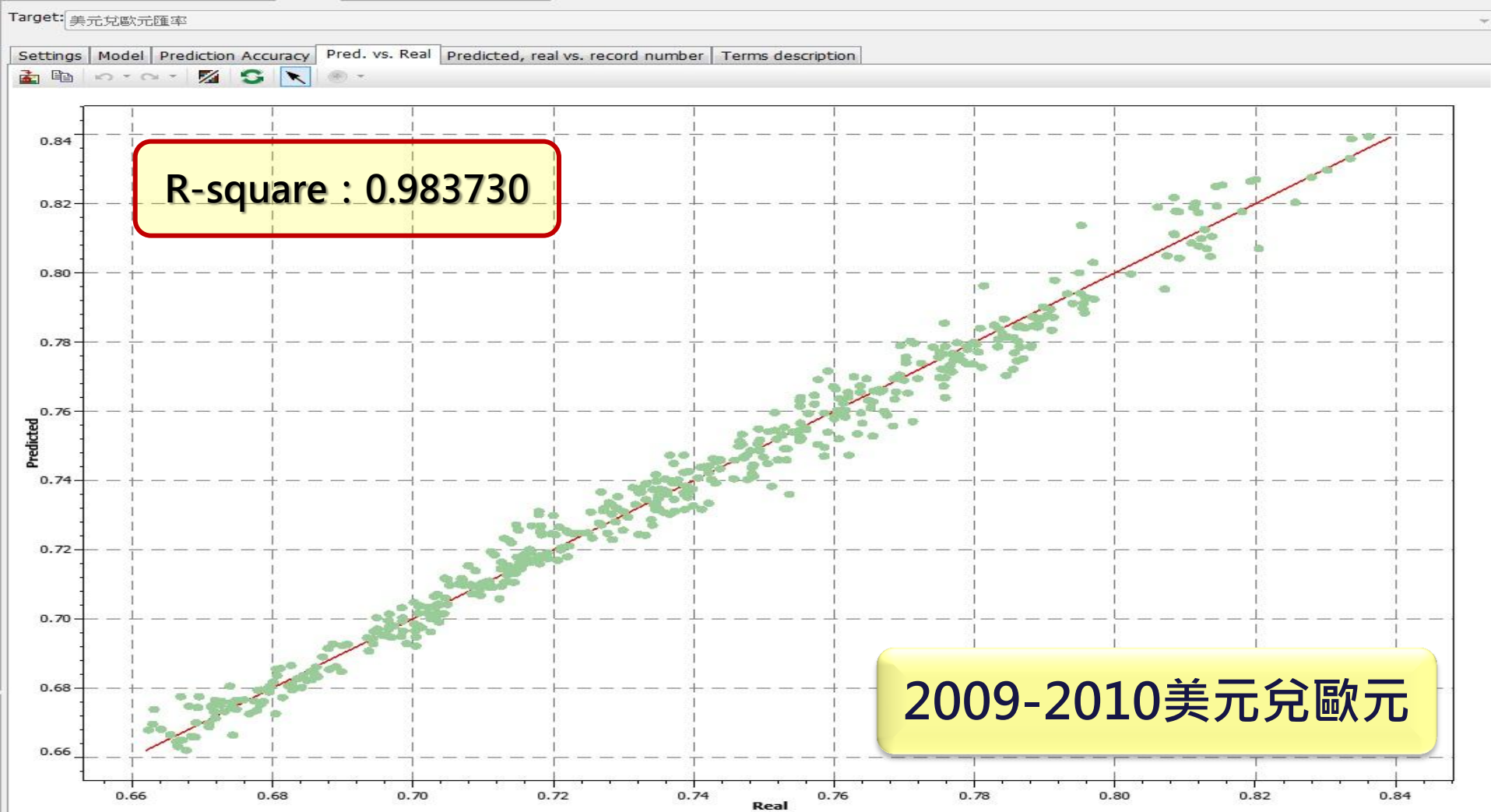
Settings Model Prediction Accuracy Pred. vs. Real Predicted, real vs. record number Terms description



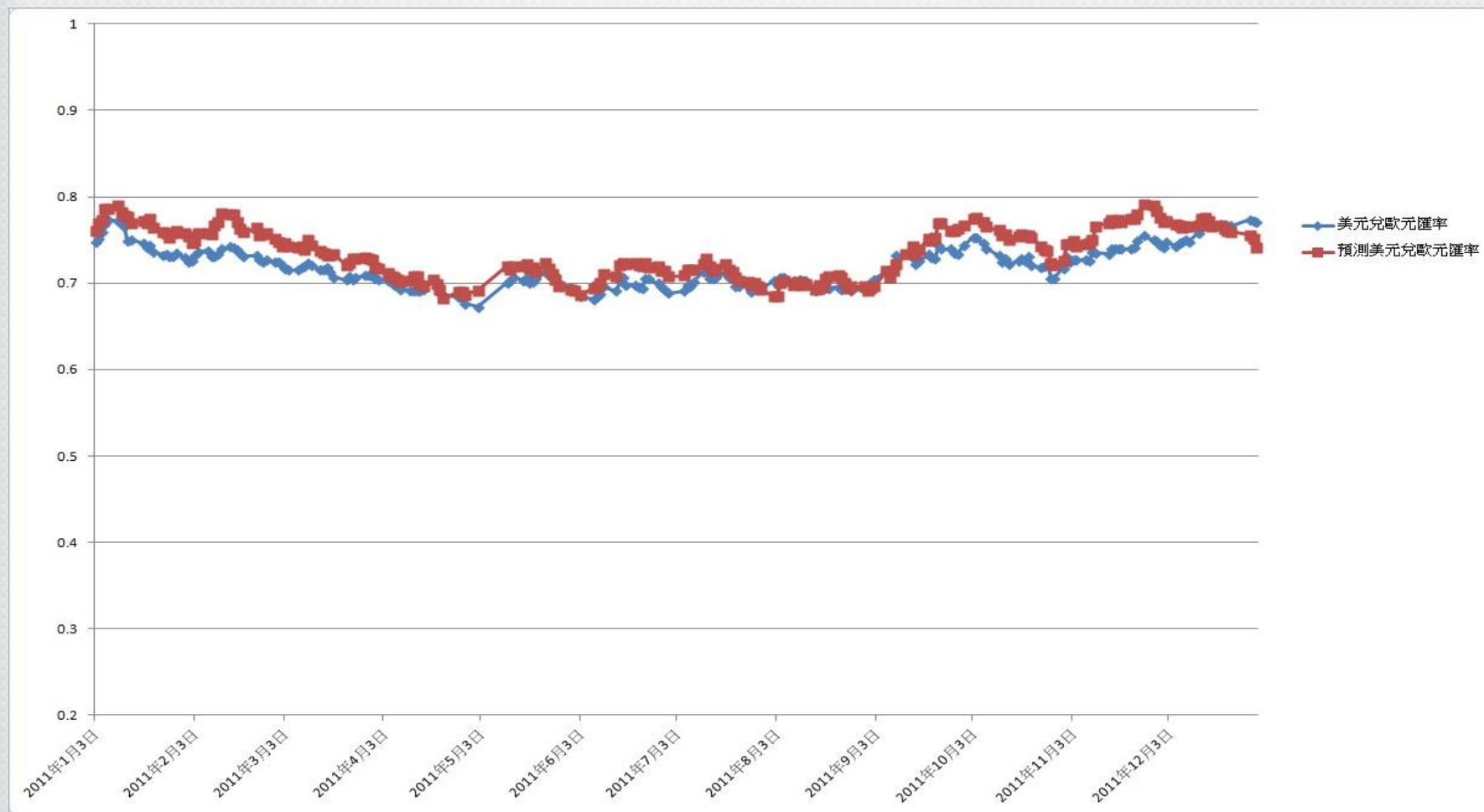
2011美元兌英鎊預測實際圖



研究方法二：美元兌歐元



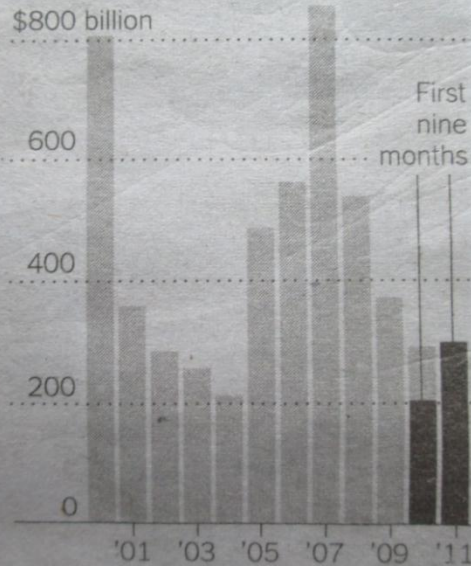
2011美元兌歐元預測實際圖



2011美元兌歐元預測

Signs of a Turnaround

Foreign direct investment in Europe rose in the first nine months of 2011 over the similar period in 2010.



Source: Organization for Economic Cooperation and Development

THE NEW YORK TIMES

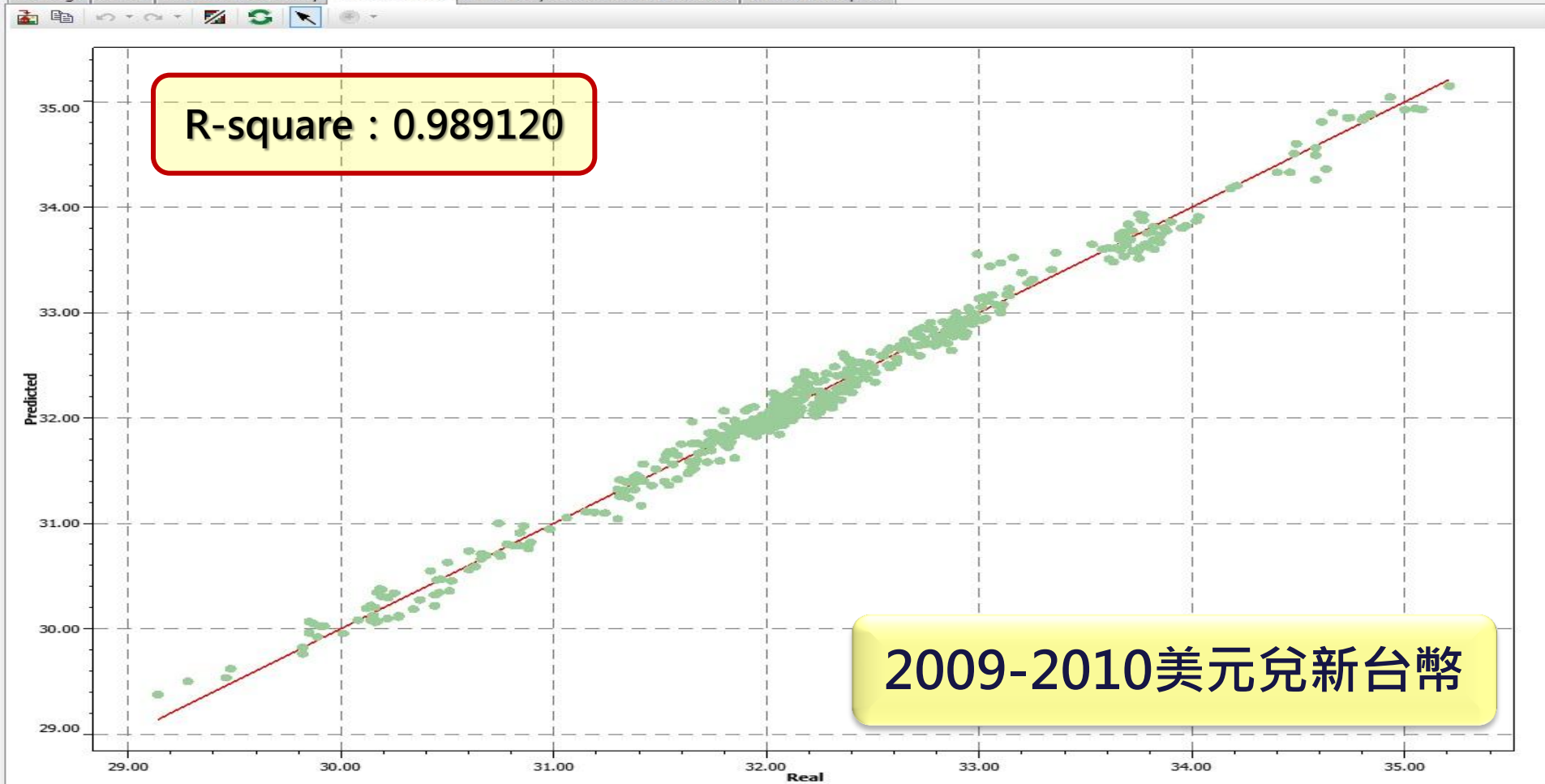
- ⊕ 2011年12月29日意大利拍賣十年期國債
- ⊕ 「歐豬」國家已換了新政府厲行節約開支，削減社會福利政策以爭取投資者的信心
- ⊕ 2012年4月The New York Times表示：外資目前正在回流歐洲



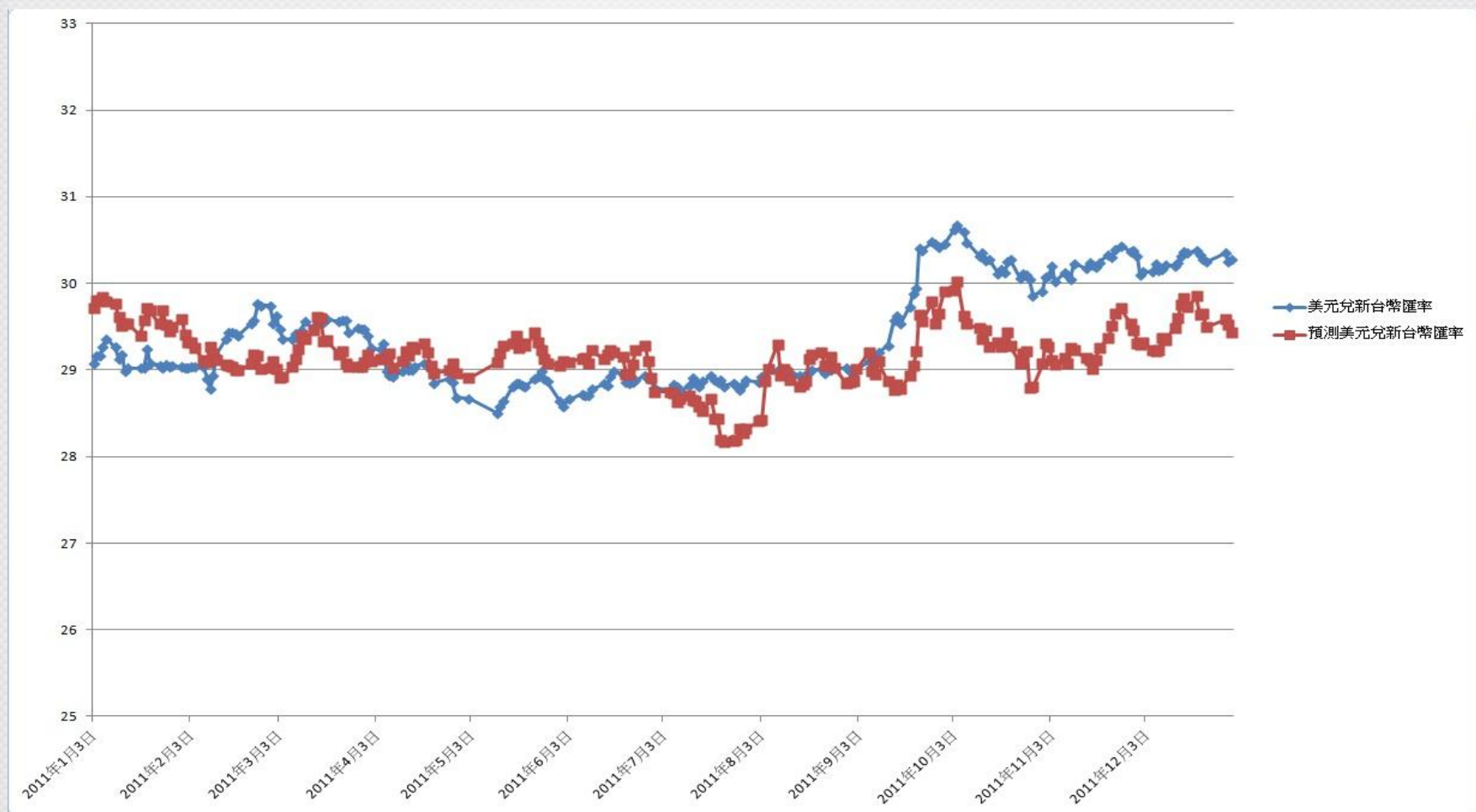
研究方法二：美元兌新台幣

Target: 美元兌新台幣匯率

Settings Model Prediction Accuracy Pred. vs. Real Predicted, real vs. record number Terms description



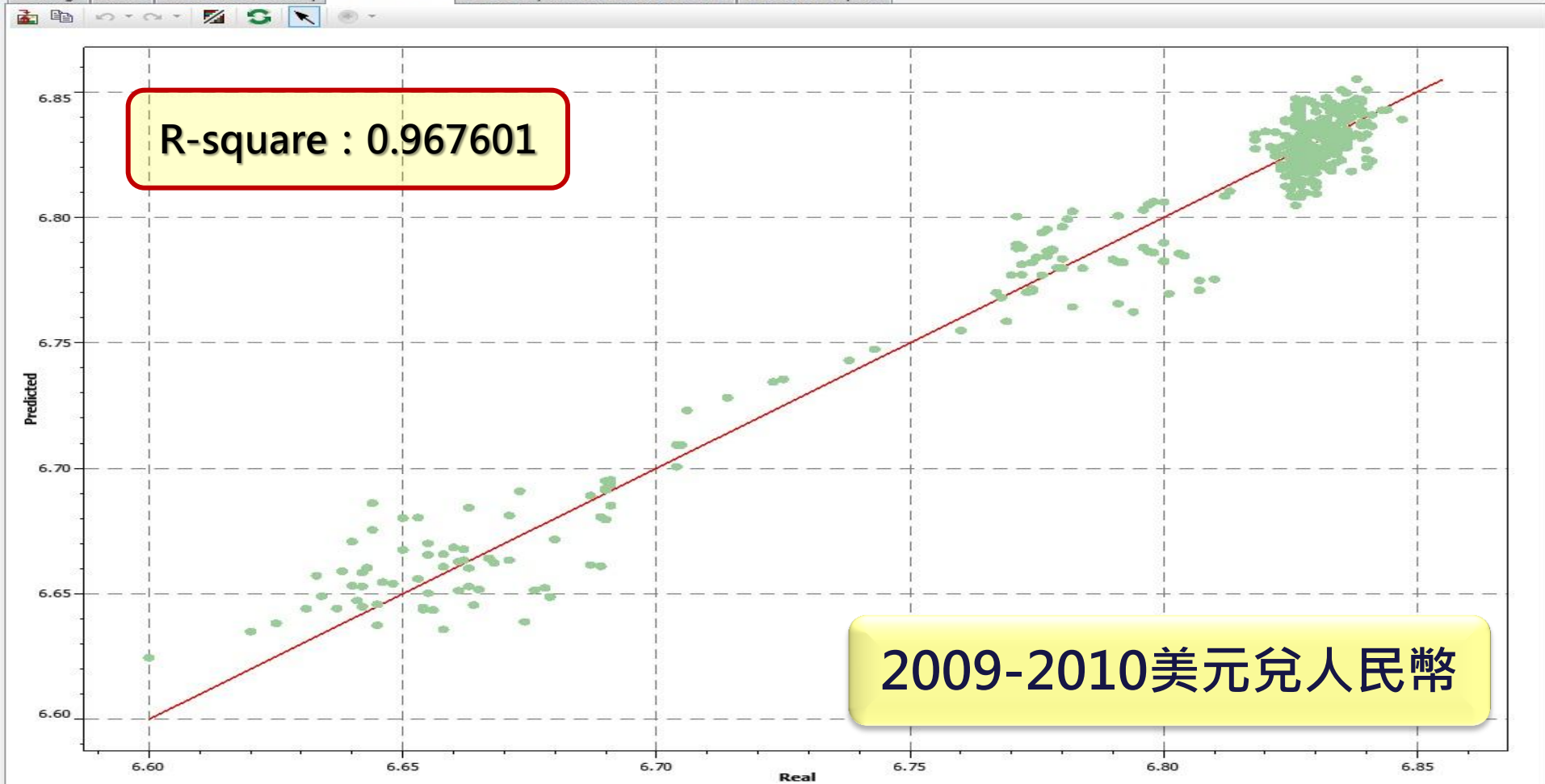
2011美元兌新台幣預測實際圖



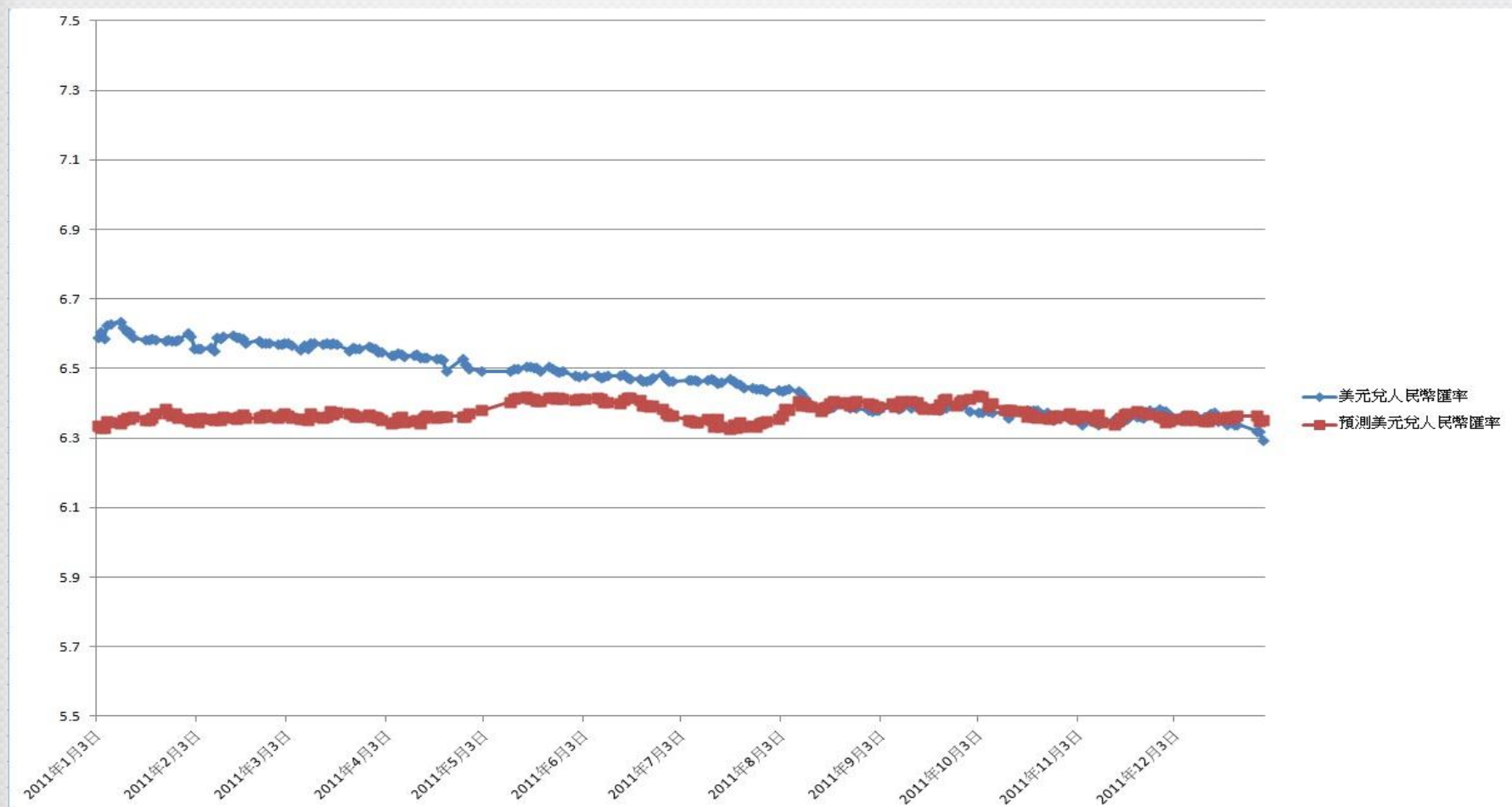
研究方法二：美元兌人民幣

Target: 美元兌人民幣匯率

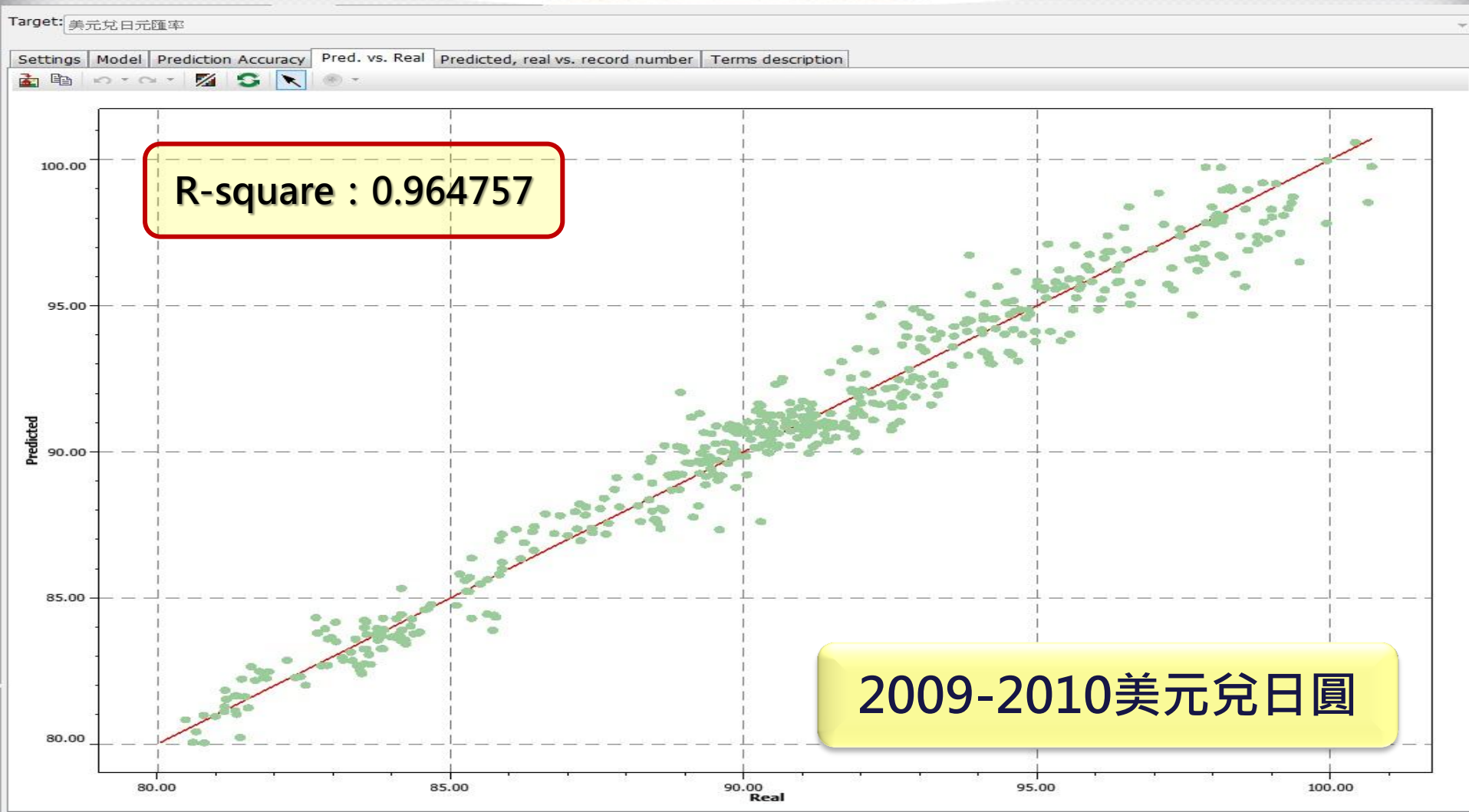
Settings Model Prediction Accuracy Pred. vs. Real Predicted, real vs. record number Terms description



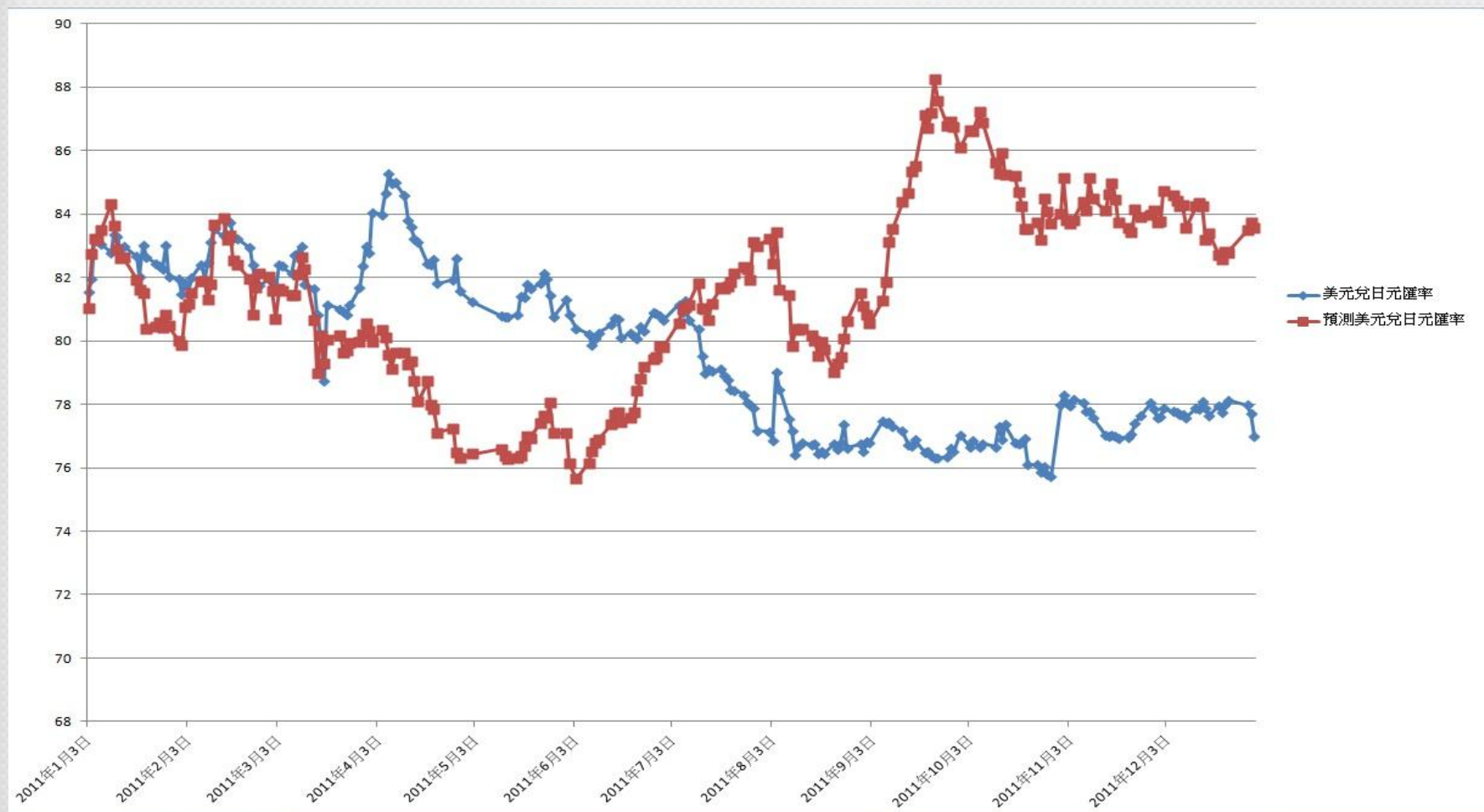
2011美元兌人民幣預測實際圖



研究方法二：美元兌日圓



2011美元兌日圓預測實際圖



研究結論



















Do we make predictions?



研究結論

- ⊕ 透過預測模組可發現影響匯率頻率較高的因子
 - 股價指數：NASDAQ、INDU
 - 原物料：Silver > gold
 - 其他：American Woodmark Corporation



美元兌台幣	美元兌人民幣	美元兌日幣	美元兌英鎊	美元兌歐元
預測模型因素	預測模型因素	預測模型因素	預測模型因素	預測模型因素
				
				
				
				
				
				
				



Corn Products International 美國伊利諾斯州玉米提煉廠



Green Mountain Coffee Roasters 美國精品咖啡業內的領導廠商



Rocky Mountain Chocolate Factor 洛磯山脈巧克力工廠



The Cheesecake Factory Incorporated 美國加州高級連鎖餐廳企業



Panera Bread Company



Overseas Shipholding Group, Inc (OSG)



Southwest Airlines Company

參考文獻

PolyAnalyst - Data Mining Software

<http://www.megaputer.com/polyanalyst.php>

維基百科·自由的百科全書

<http://zh.wikipedia.org/zh-tw/Wikipedia:%E9%A6%96%E9%A1%B5>

台灣證券交易所 國際股市年報

http://www.twse.com.tw/ch/statistics/statistics_list.php?tm=07&stm=032

Economic Research

<http://research.stlouisfed.org/fred2/series/CPIAUCSL/downloaddata>

YCharts

<http://ycharts.com/>

談弱勢美元

<http://www.rusrule.com/rusrule/?p=836>

美QE2政策的效果與副作用

<http://www.npf.org.tw/post/1/8379>

中國評論新聞：美元崩潰、全球蕭條與美國解體？

<http://www.chinareviewnews.com/doc/1014/9/7/6/101497617.html?coluid=7&kindid=0&docid=101497617>





THE END

銘謝

東吳大學資訊管理學系趙景明教授
東吳大學財務精算學系助教Roy