ÚOHS ÚŘAD PRO OCHRANU HOSPODÁŘSKÉ SOUTĚŽE

Retail fuel data analysis

Zbyšek Martoch
Marek Smysl

Structure of presentation

I. Data Set
II. Transformation into Relational Database
III. Transactions analysis
IV. Price analysis
V. Geographical Visualisation
VI. Further Possible Development
VII. Possible Utilization

I. Data Set

- Source of data

  CCS Česká společnost pro platební karty, s. r. o.

  (29th April 2013)

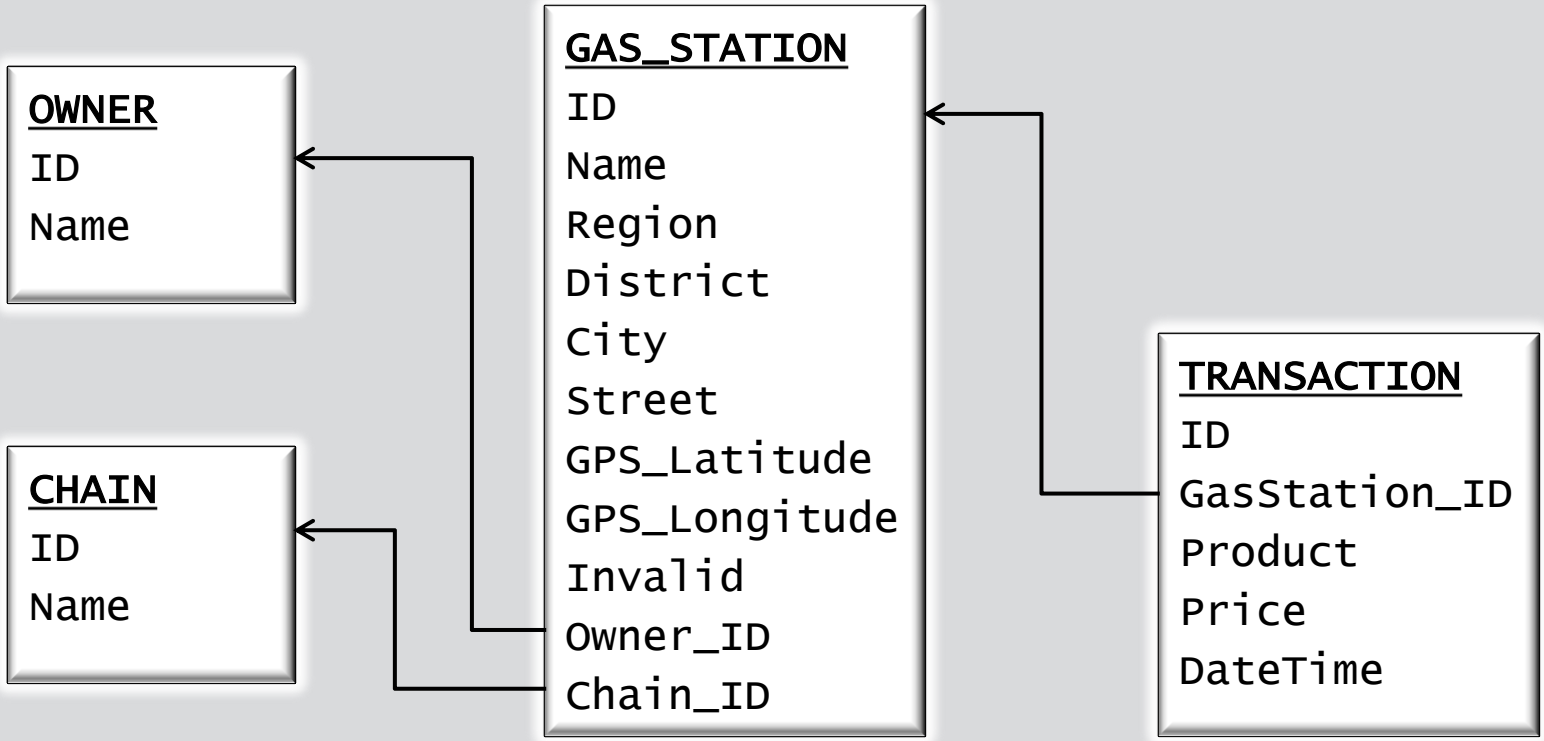
CCS is the largest company offering fuel (credit) cards for business clients in the Czech Republic

Structure of Data

- Period: 1st January 2008 till 31st March 2013
- Data Format: CSV
  - Type of fuel
  - Price
  - Time and date of transaction
  - Place of transaction (verbal designation)
  - Place of transaction (GPS coordinates)

II. Flat data transformation into relational database

- Huge database – over 6 GB
- Data handled by SQL scripting



Relevant part of database scheme

- Hosted on OHE server, accessible within LAN
- Ready for periodical uploading of fresh data in future
- Basic statistics and reports created directly by SQL
- Advanced data processing performed by Excel, SPSS and Gretl

III. Transactions analysis

Initial conditions

- over 50 mil. payment transactions
- 10 different vague identified products in dataset
- only distinction between petrol and diesel tagged

results

- “TOP 2” products cover ~ 95 % of transactions
- “TOP 4” products cover ~ 98 % of transactions
- Detected reliably: Natural95, standard Diesel, Premium Natural (Natural 98, brand specific fuels), Premium Diesel
- other products < 2% → dropped
- Detected periodic variation in transaction time series with significant week and month period.

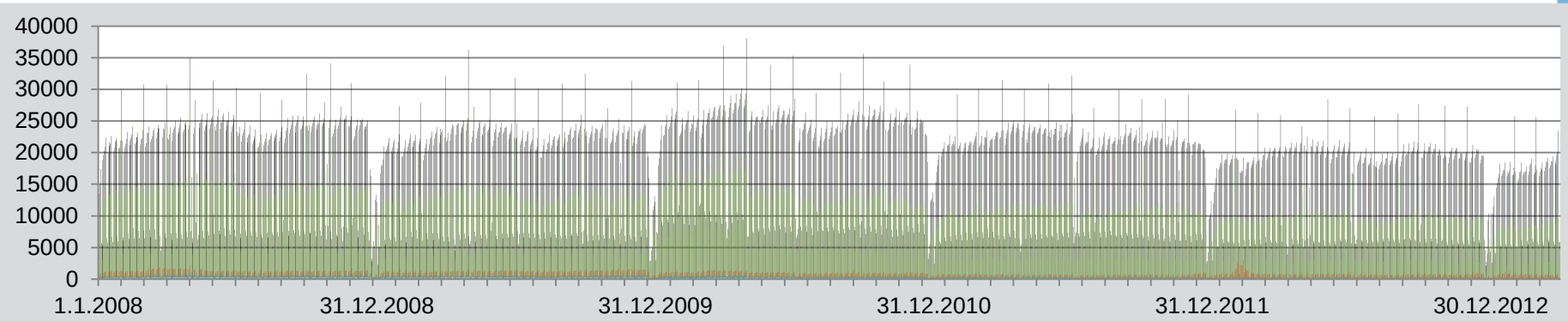
Considering that the dataset contains only numbers of CCS transactions, the trend analysis of number of transactions has only a minor importance. Numbers of transactions were preferably used to estimate the ratio of petrol stations turnovers.

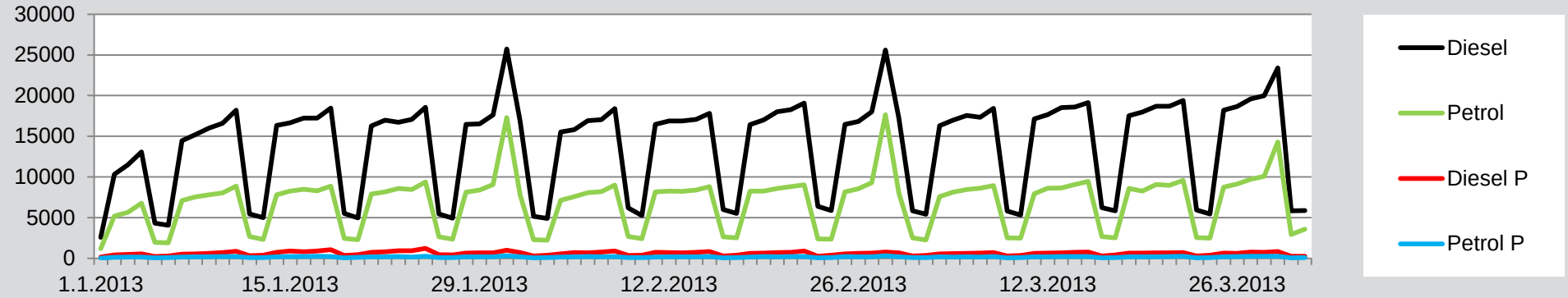
- Number of Petrol Stations in Data Set (April 2013)

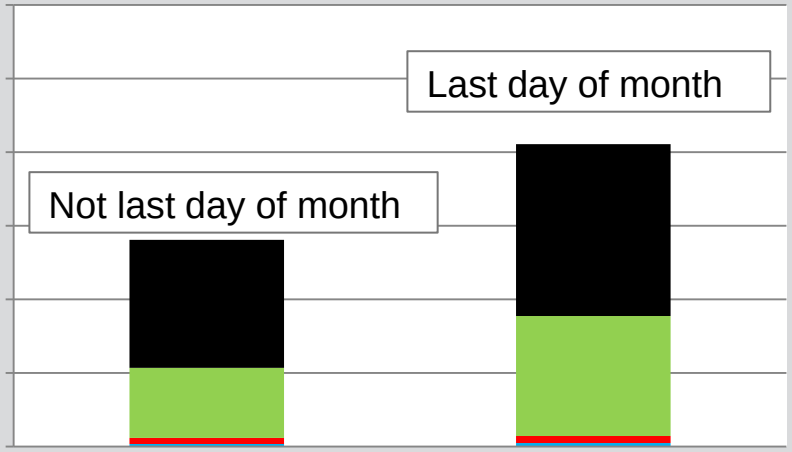
	Company	Number
1	BENZINA, s. r. o.	334
2	OMV ČESKÁ REPUBLIKA,S.R.O.	205
3	ČEPRO A.S.	195
4	SHELL CZECH REPUBLIC, A.S.	159
5	ENI ČESKÁ REPUBLIKA S.R.O.	133
6	PAP OIL ČERPACÍ STANICE S.R.O.	132
7	ROBIN OIL , S.R.O.	73
8	LUKOIL CZECH REPUBLIC S.R.O.	44
9	KM-PRONA,A.S.	33
10	TANK ONO SPOL. S R.O.	33
11	SLOVNAFT ČESKÁ REPUBLIKA S.R.O	27
12	UNICORN-ČERPACÍ STANICE,A.S.	24
13	AHOLD CZECH REPUBLIK-ALBERT HM	23
14	SOUKROMÉ ČS-FRENČÍZY SHELL	23
15	SILMET HP,A.S.	21
16	TESCO STORES ČR A.S.	18

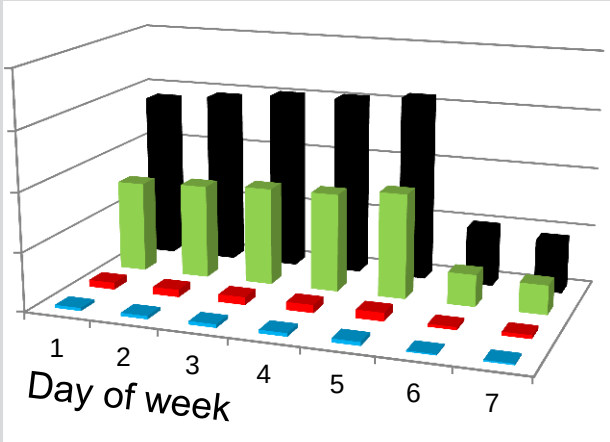
	Company	Number
17	PASOIL S.R.O.	17
18	GLOBUS ČR, k.s.	15
19	SOUKROMÉ ČS-FRENČÍZY OMV	15
20	KONTAKT-SLUŽBY MOTORISTŮM	14
21	ARMEX OIL, s. r. o.	13
22	MAKRO CASH & CARRY ČR S.R.O.	13
23	STANISLAV ŠEFL - MEDOS	10
24	ČERPACÍ STANICE LPG	8
25	HUNSGAS S.R.O.	7
26	SOUKROMÉ ČS-FRENČÍZY BENZINA	7
27	GEDAL A.S.	4
28	SWISS TANK S.R.O.	4
31	KARIMPEX, A.S.	1
32	UNIXAN S.R.O.	1
33	SOUKROMÉ ČERPACÍ STANICE	781
	T o t a l	**2387**

Daily numbers of transactions by type of fuel

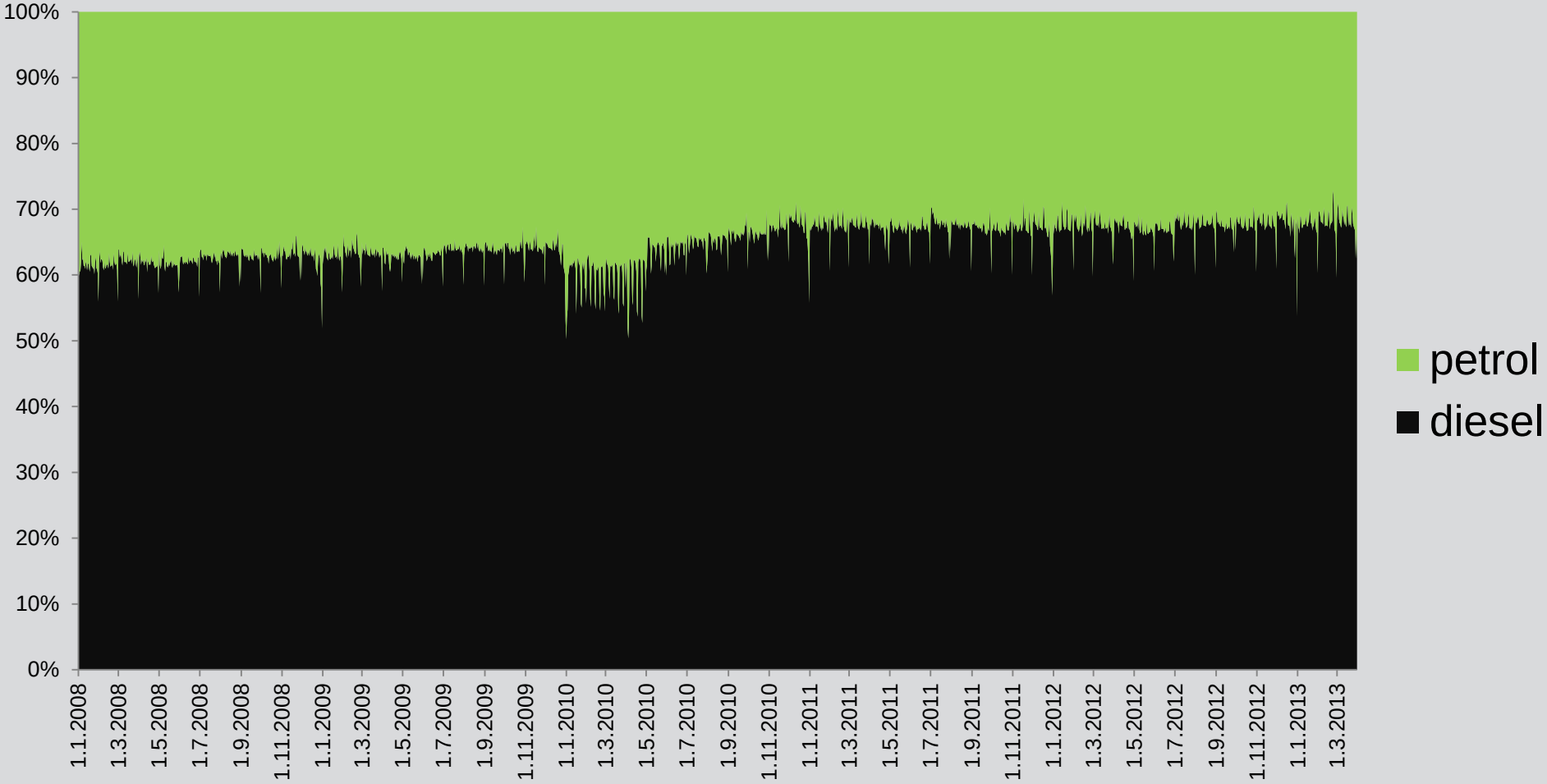








Proportion of Transaction – Diesel vs. Petrol



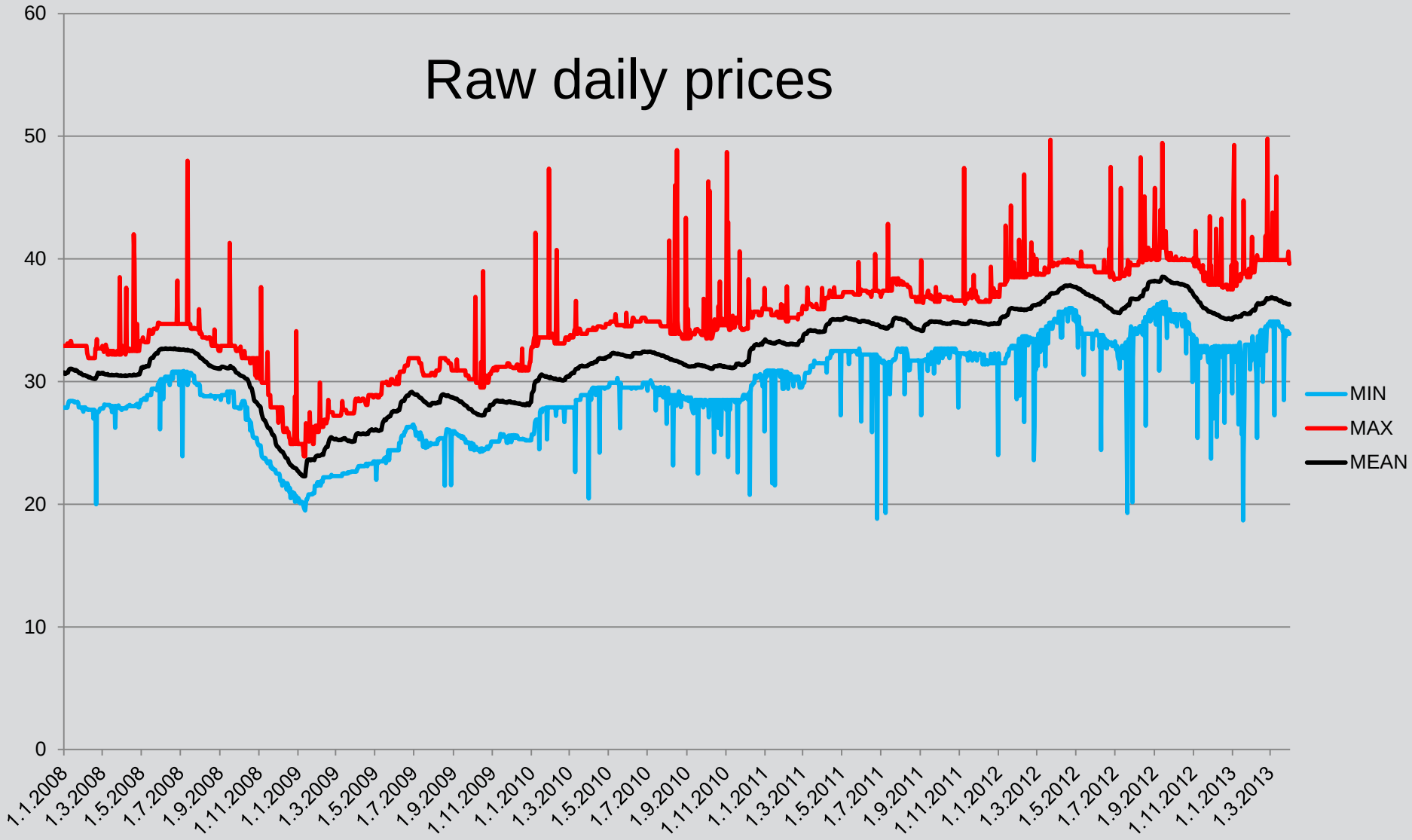
IV. Price analysis

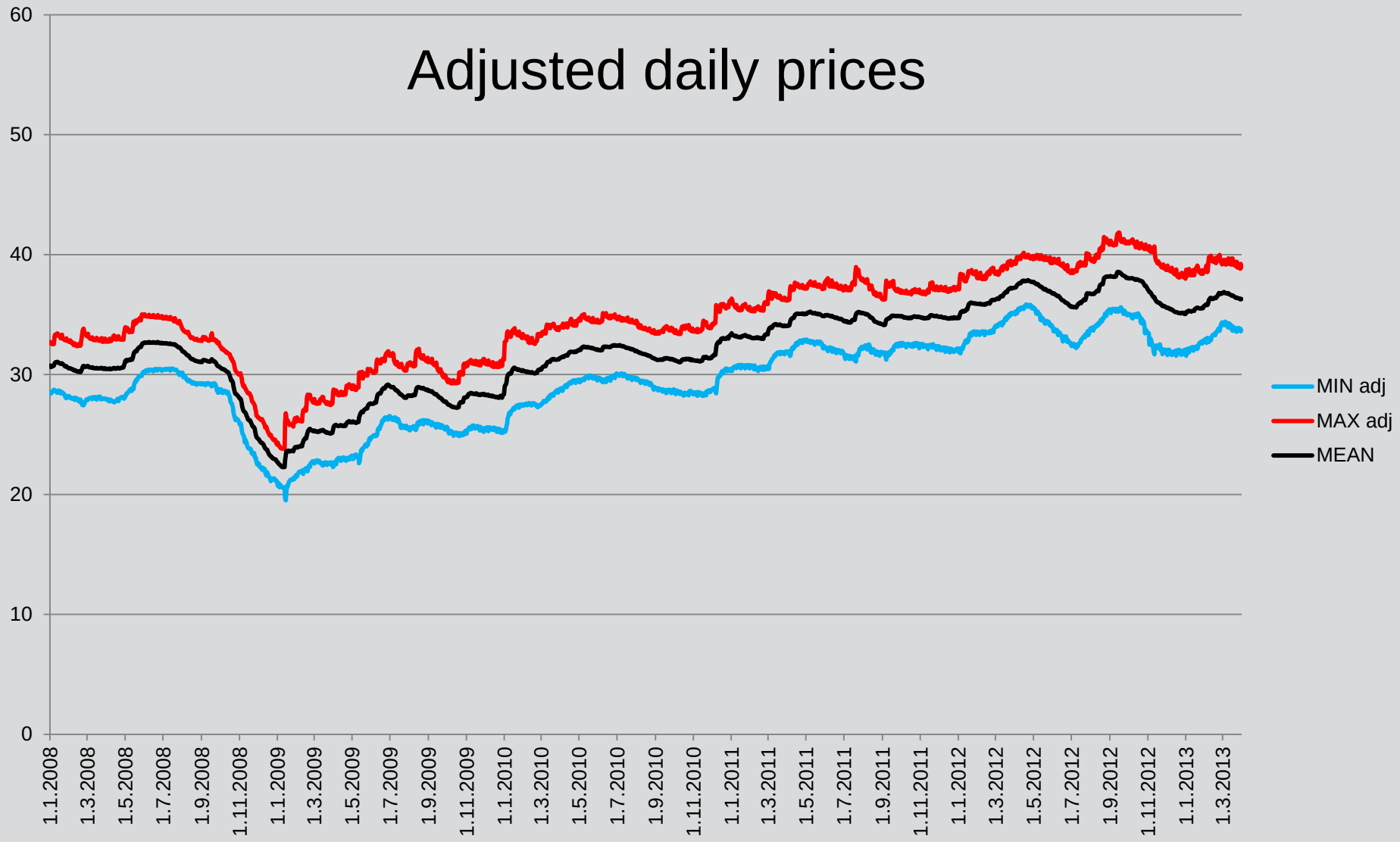
approach

- Detected less than 0.1 % outliers (evidently errors) and dropped
- All daily prices were normalized into within the interval <0;1>
- 2 top products were considered ~ 95 % of data

results

- All prices are highly correlated – The part of wholesalers prices is necessary to remove from retailers prices.
- The motorways (D1, D2, D5 and D11) petrol stations sell at significantly higher prices. D3 and D8 petrol stations do not.
- Uneven prices changes in particular regions. North Moravia is getting more expensive over our period than the others.
- Several sharp price increases were detected, and some of them were explained.


Raw daily prices
60
50
40
30
20
10
0
1.1.2008
1.3.2008
1.5.2008
1.7.2008
1.9.2008
1.11.2008
1.1.2009
1.3.2009
1.5.2009
1.7.2009
1.9.2009
1.11.2009
1.1.2010
1.3.2010
1.5.2010
1.7.2010
1.9.2010
1.11.2010
1.1.2011
1.3.2011
1.5.2011
1.7.2011
1.9.2011
1.11.2011
1.1.2012
1.3.2012
1.5.2012
1.7.2012
1.9.2012
1.11.2012
1.1.2013
1.3.2013
MIN
MAX
MEAN


Adjusted daily prices
60
50
40
30
20
10
0
1.1.2008
1.3.2008
1.5.2008
1.7.2008
1.9.2008
1.11.2008
1.1.2009
1.3.2009
1.5.2009
1.7.2009
1.9.2009
1.11.2009
1.1.2010
1.3.2010
1.5.2010
1.7.2010
1.9.2010
1.11.2010
1.1.2011
1.3.2011
1.5.2011
1.7.2011
1.9.2011
1.11.2011
1.1.2012
1.3.2012
1.5.2012
1.7.2012
1.9.2012
1.11.2012
1.1.2013
1.3.2013
MIN adj
MAX adj
MEAN

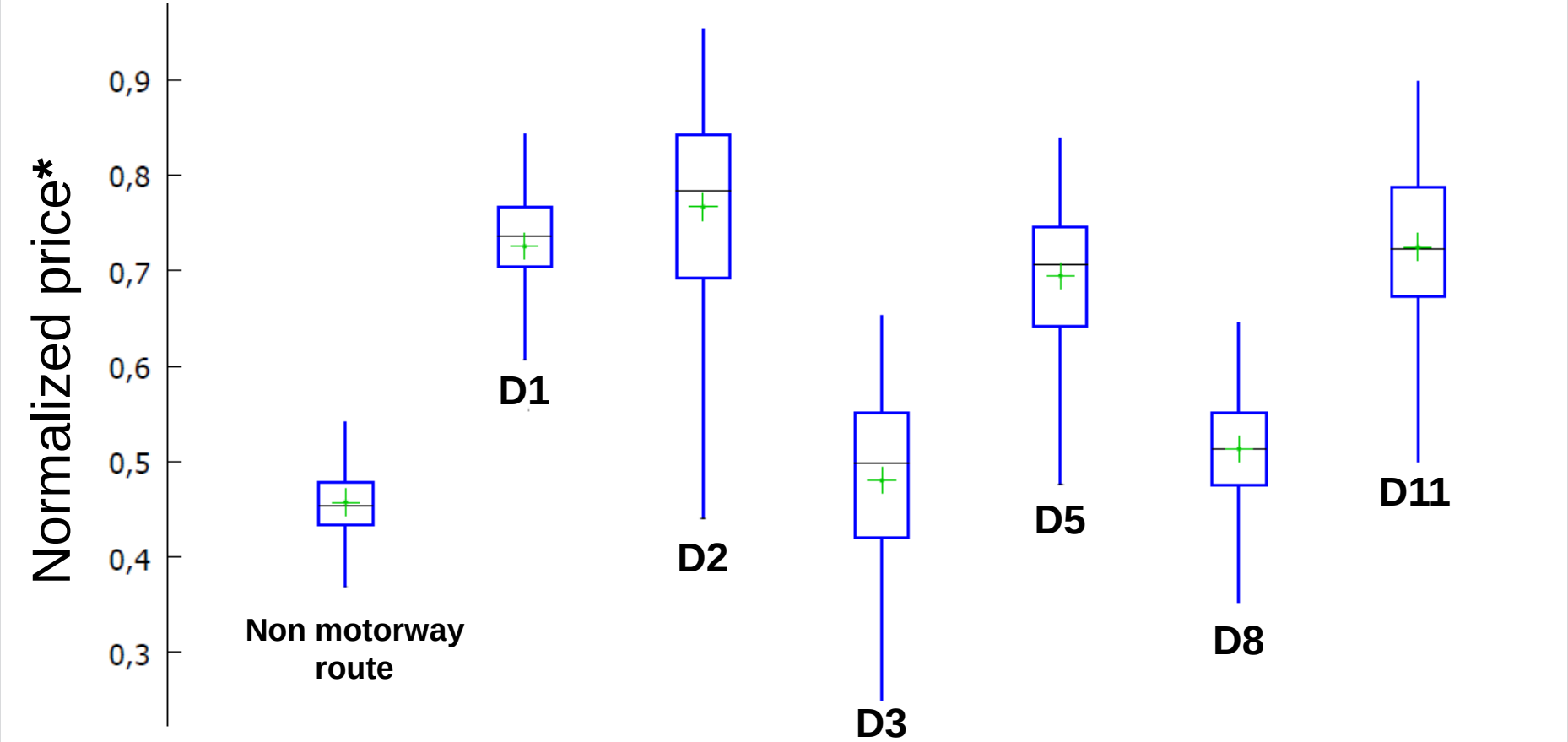
- Price Characteristics [CZK/litre]

	DieselP	*Diesel*	*Petrol*	*PetrolP*
Mean	34.92	32.01	32.32	34.53
Median	35.09	32.53	32.41	34.76
Standard deviation	4.08	3.81	3.71	3.74
Variance	16.67	14.50	13.75	13.98
Minimum	25.81	24.08	22.40	24.11
Maximum	40.48	37.30	38.52	40.59

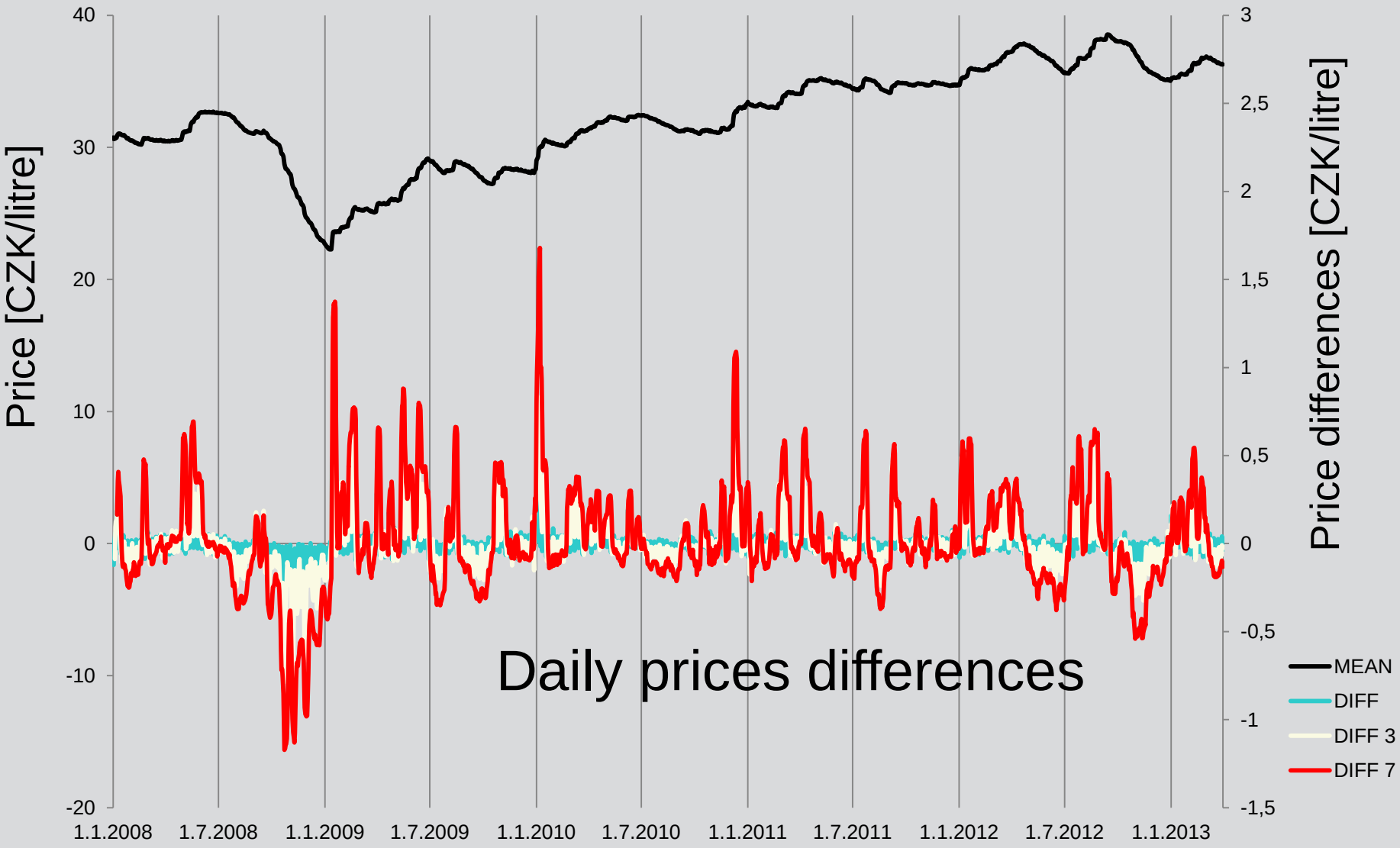
- Price difference between Standard and Premium Fuel

Diesel vs. DieselP	*Petrol vs. PetrolP*
9.10%	6.84%

Comparison of prices at motorway and off-motorway petrol stations



* Average daily price at each gas station was normalized into interval <0;1> based on min and max prices which occurred on the particular day.


Daily prices differences
Price [CZK/litre]
Price differences [CZK/litre]
40
30
20
10
0
-10
-20
3
2,5
2
1,5
1
0,5
0
-0,5
-1
-1,5
1.1.2008
1.7.2008
1.1.2009
1.7.2009
1.1.2010
1.7.2010
1.1.2011
1.7.2011
1.1.2012
1.7.2012
1.1.2013
MEAN
DIFF
DIFF 3
DIFF 7

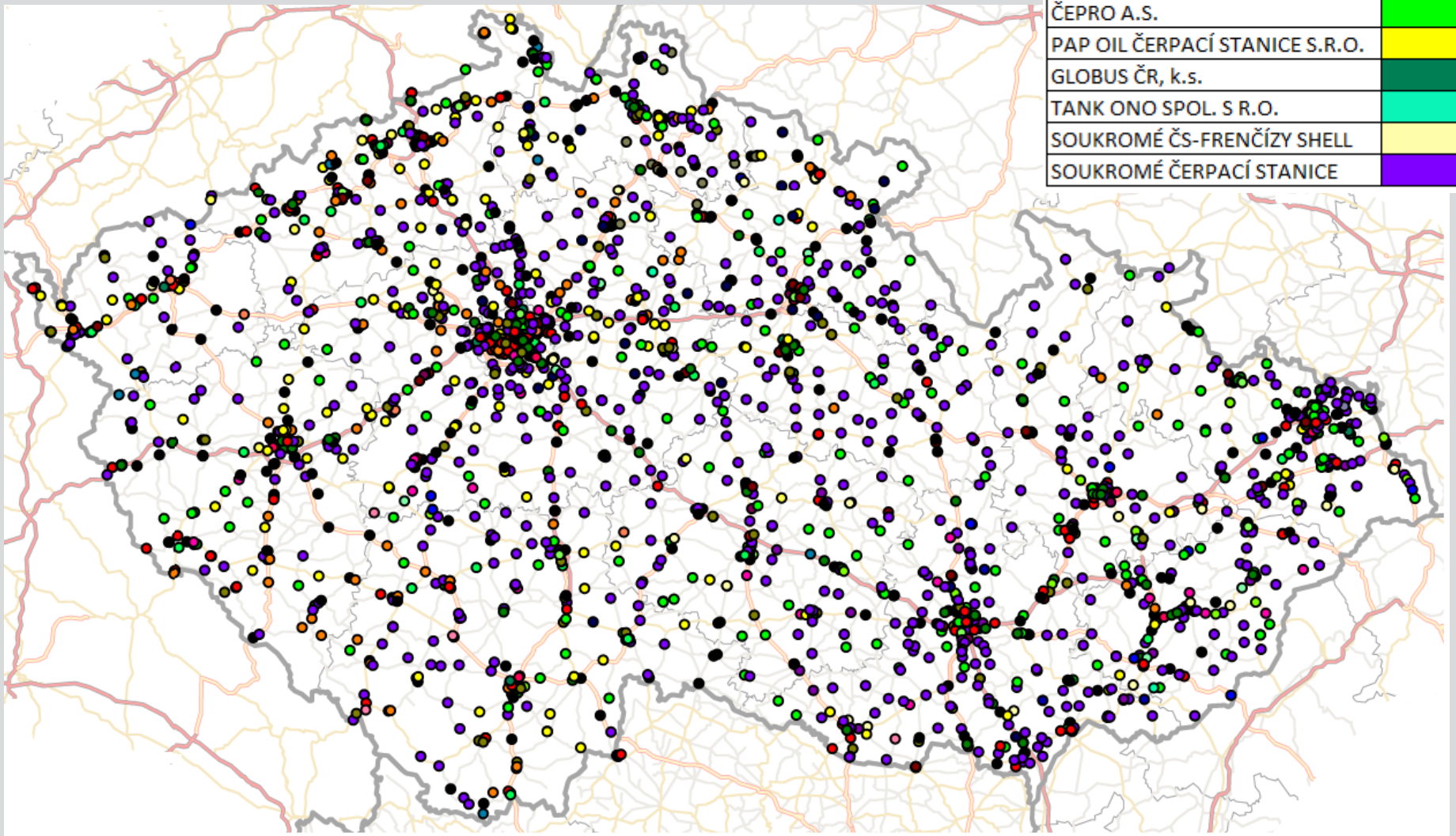
V. Geographical Visualisation

- 98 % of petrol stations were localized (GPS position in dataset available)
- only the last owner of the petrol station is known, owners history is missing
- „Motorway factor“ was entered in proper dummy variables
- Special software for geographical analysis of dataset was developed

ÚOHS ÚŘAD PRO OCHRANU HOSPODÁŘSKÉ SOUTĚŽE

Petrol stations positions

BENZINA, s.r.o.	
LUKOIL CZECH REPUBLIC S.R.O.	
OMV ČESKÁ REPUBLIKA,S.R.O.	
SHELL CZECH REPUBLIC, A.S.	
ENI ČESKÁ REPUBLIKA S.R.O.	
ROBIN OIL , S.R.O.	
ČEPRO A.S.	
PAP OIL ČERPACÍ STANICE S.R.O.	
GLOBUS ČR, k.s.	
TANK ONO SPOL. S R.O.	
SOUKROMÉ ČS-FRENČÍZY SHELL	
SOUKROMÉ ČERPACÍ STANICE	



VI. Further Possible Development

- Evaluation of traffic density and price behaviour of petrol stations according to the type of a road (based on data from *Road and Motorway Directorate of the Czech Republic*)
- Measuring and Comparing of retail margins (upstream and downstream market data necessary)
- Evaluation of the impact of advanced deposit (20 mil. CZK for all retailers; 1st November 2013)

VII. Utilizations

- Mergers
- Cartels
- Abuse of dominance
- Shock analysis

Thank you for your attention