



# Container transport chains and codification of commodities

Peter Smeets

Senior statistical researcher

**Statistics Netherlands**

Working Party on Transport Trends and Economics  
*Group of Experts on Hinterland Connections of Seaports*

Geneva, 28 October 2008



# Overview of the presentation



Statistics Netherlands

- Users' Needs
- Data sources
- A. Transport chains
  - Linking sources (problems en possibilities)
  - Examples
- B. Commodities
  - Free text
  - Coding system
  - Validation
  - Provisional results
- Future developments



# Disclaimer



Statistics Netherlands

- "All data are provisional and may not be quoted or published"
- "The content of this presentation represents the personal views of the author and need not necessarily coincide with those of Statistics Netherlands "
- It is a pilot study and incorporation into regular program is to be decided on



# Users' needs



Statistics Netherlands

Key users of statistical data (freight transport) rank **missing information:**

1. Content of containers
2. Transport chains (hinterland connections)
3. Transit with transshipment

This information is **important** for

- infrastructural planning
- Expanding activities/ marketing
- Modal shift programs



# Data sources



Statistics Netherlands

- Detailed information available from incoming maritime transport reported to CUSTOMS (Sagitta system, pre-arrival information)
  - Container identification numbers
  - Commodity description (free text field)
- Railway companies
  - Container identification numbers
- Barge information and communication system (BICS) (edifact based system)
  - Container identification numbers



# A. Transport chains





# Pilot project part A



Statistics Netherlands

- Intention to link incoming maritime container transport with:
  - Inland waterways (BICS)
  - Railways statistics

linking based on: container number

In order to model the MODAL SPLIT



# A1. Railway statistics



Statistics Netherlands

- Project steps
- Develop method to link information based on the container numbers (micro approach)
- Use micro data to develop a model for transport links (macro approach)
- Produce output: transit with transshipment tables
- Advise whether standard output on regular base is feasible.





# Current status



Statistics Netherlands

- First steps in linking micro data (maritime-railways)
- Problems with data quality: almost 50% of the container numbers misses the last digit (control digit).
- Customs data April 2007 linked to Railway data April and May 2007.
- Linking on container number and nearest date.



# Results (1)



Statistics Netherlands

- 25% of the “railway” containers loaded in a maritime area could be linked to the “maritime” containers
- Problems with short distance traffic:  
Cycle is too short
- Empty containers:  
98% intra EU-transport (UK, IE)
- First analyses concentrate on  
DEEP-SEA full container loads (FCL)



# Results (2)



Statistics Netherlands

Number of containers (FCL) trans shipped via NL (DEEP SEA-RAIL) by country of loading and country of unloading, April 2007

Top 10 country of loading (SEA)	Country of unloading (RAIL)										
	CH	DE	AT	BE	PL	SK	HU	IT	NL	FR	total
CN	462	310	419	156	45		38	31	6		1467
SG	198	69	55	217	6	4	5	2	4		560
KR	72	172	9	12	99	135	19		1		519
JP	31	283	26				6	8			354
MY	82	169	28				1				280
US	106	42	30	1	35		21	6			241
HK	77	18	52	11	4	4	21	5	3		195
IN	68	7	72	3			1			1	152
TW	59	9	19	1		31	7		1		127
ZA	20	3	52		8			26			109
<b>Total</b>	<b>1175</b>	<b>1082</b>	<b>762</b>	<b>401</b>	<b>197</b>	<b>174</b>	<b>119</b>	<b>78</b>	<b>15</b>	<b>1</b>	<b>4004</b>



# A2 Inland waterways (IWW)



Statistics Netherlands

- Electronic Data from Barge Information and Communication system (BICS) are used (April and May 2007)
- Data linked by container number
- Problem: IWW journeys can reported more than once: result double counting



# Example: Inland waterways

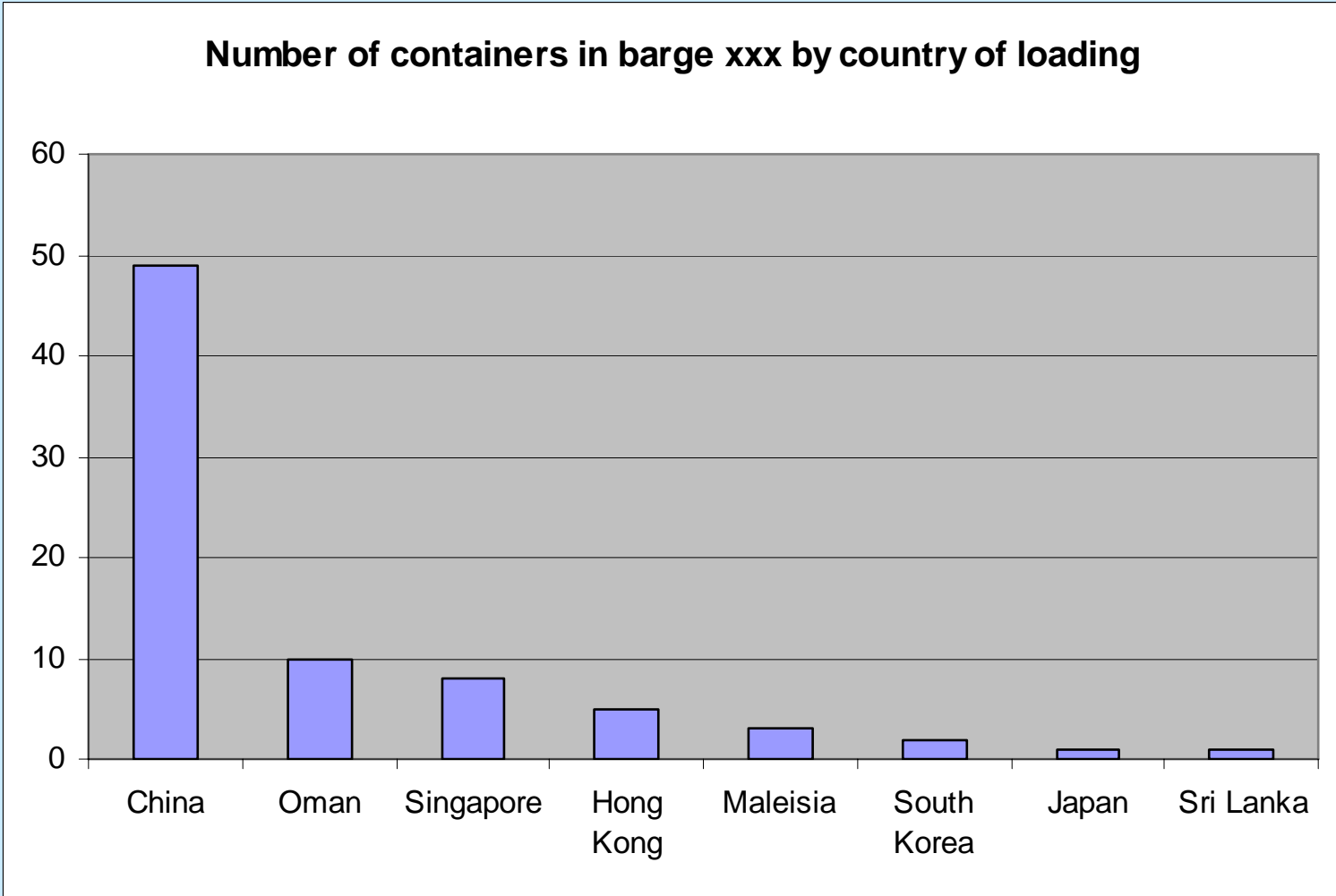


Statistics Netherlands

- The ship xxxxx sails on 23-4-2007 from **Rotterdam (NL)** to **Duisburg (DE)**.
- It carries **114 containers**.
- Of which 20 empty and **94 loaded**.
- **79** of the loaded containers can be linked with maritime data (Customs data) by containernumber.
- **58 containers** are 20 ft and **21 containers** are 40 ft (or more)
- Total number of TEU (Twenty feet Equivalent Units) :  
 **$58 + 21 * 2 = 100$  TEU**
- Some containers have **more than 1 shipment** (total 101 shipments)
- Total weight of the goods is **1437 tonnes**
- **8 different ship brokers** provided the information to the customs
- The containers were carried by **15 different sea vessels** that called Rotterdam between **1 and 22 April 2007**.

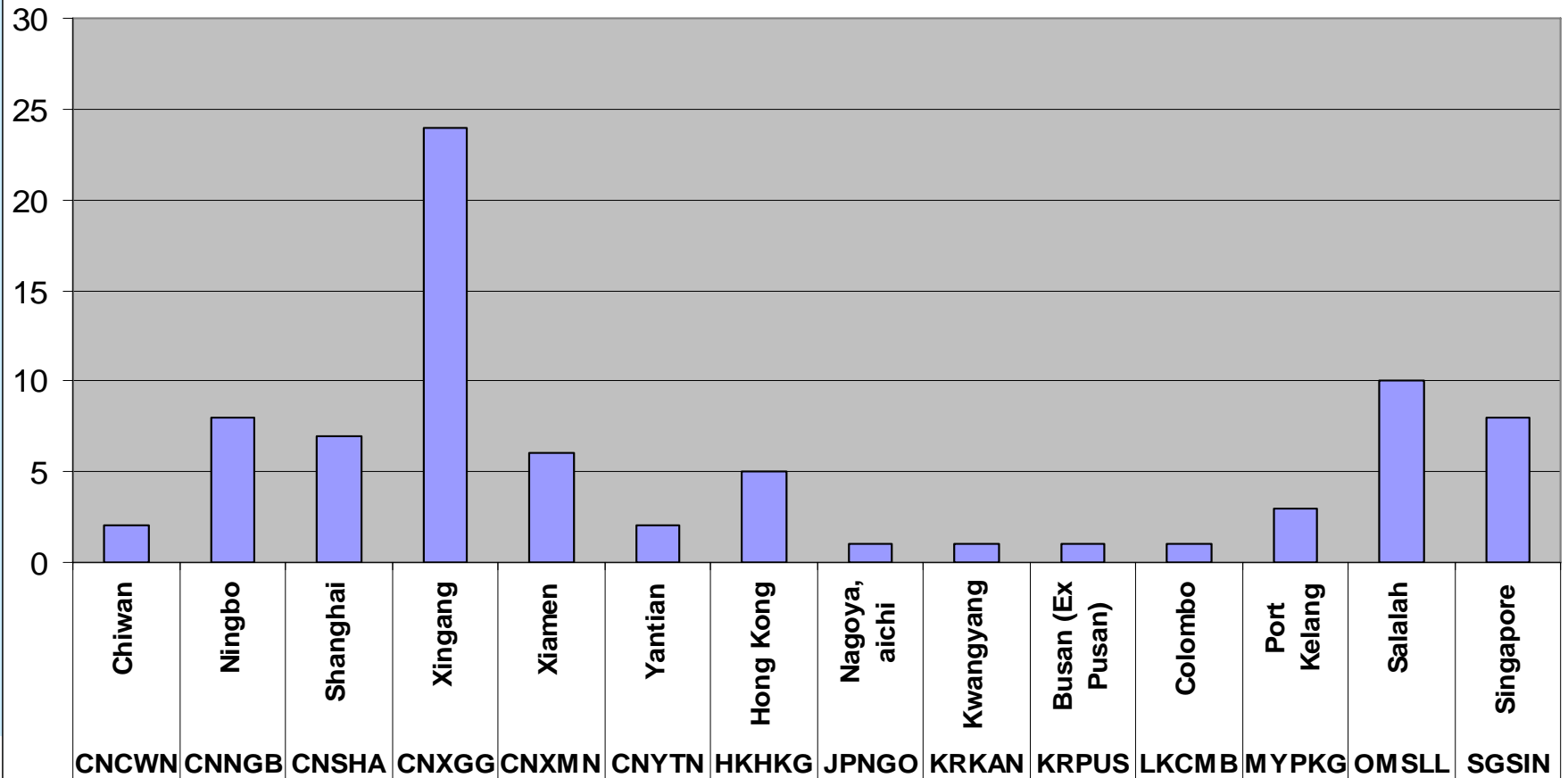


### Number of containers in barge xxx by country of loading





## Number of containers on barge xxxx via Rotterdam to Duisburg by place of loading





# Future work



Statistics Netherlands

- Further development of linking algorithms
- Eliminating “false” linking
- Analysing representativity of linked data
- Setting up a model tot estimate the modal split





## B. Commodities

**Most important  
user question  
in freight transport:**

**What is inside the containers?**



## Facts:

- Since the transit statistics have been stopped in the mid '90ties, there is no information about the content of containers that are entering or leaving the Netherlands.
- Trade statistics only describe import and export and not transit with transshipment.



# Why now (1)



Statistics Netherlands

- **The container transport shows a yearly increase of 10% worldwide since 1995.**
- **25-30% of the goods unloaded in Rotterdam is containerised**
- **Further growth is expected.:**
  - **on 5 September 2008 a new terminal in Rotterdam is opened with a capacity of 2,3 mln containers (TEU) per year.**
  - **Extension of the port of Rotterdam is started. Partly used for container terminals.**
  - **Exploitation of the new railway line “the Betuwelijn” heavily leans on container transport**



## Why now (2)



Statistics Netherlands

- Modal choice of the hinterland transport depends partly on the type of commodity.
- Political choices can better be founded if there is more knowledge about the commodities in containers. (“promotion of modal shift”)



What do we know about the content of  
containers?

Nothing

!!

(or by coincidence...)







Beachcombing after an accident with Container ship Napoli, Jan 2007

40.000 bibles, printed in the Netherlands on its way to South-Africa





Statistics Netherlands



# SHIPWRECK

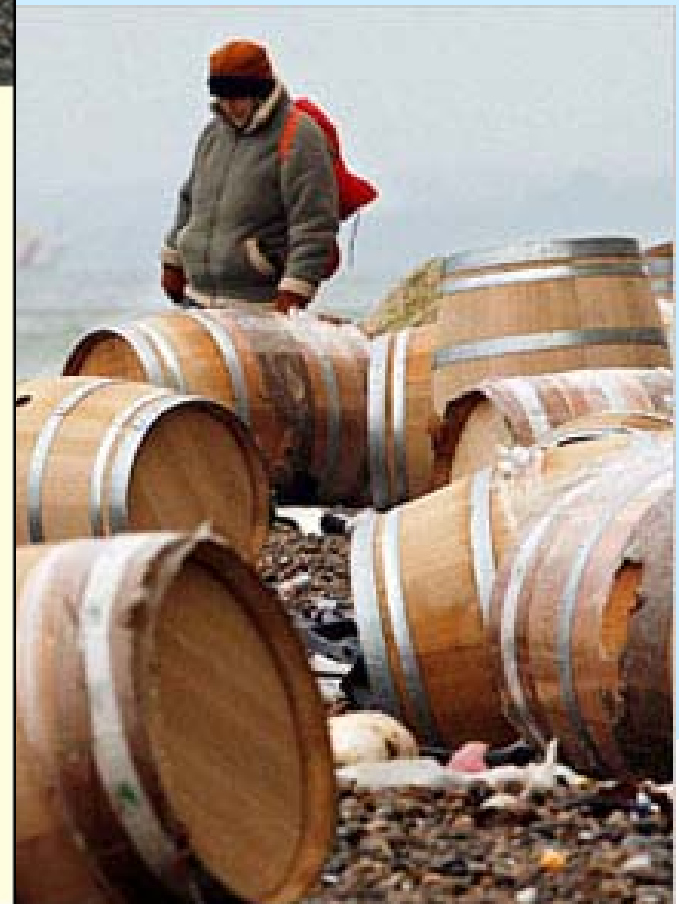
This 10 year old Cider Brandy has been finished in Allier oak casks retrieved from the stricken MSC Napoli which sank off Branscombe in Devon in January 2007.

The barrels, which were protected from the sea by bibles written in Zulu, were en-route to South Africa to age some of their finest wines. The oak produces a sophisticated and subtle finish which enhances the flavours in our brandy.

50cl

Single Cask

%Vol





# Why *automatic* coding of commodities ?



Statistics Netherlands

- Data is or will be available in free-text format (electronically) ( CUSTOMS pre-arrival system Sagitta for maritime and in soon also for aviation )
- Too much data to codify manually
- Knowledge of data processing workers can be stored in rules and conversion tables (from “heads to tables”)
- **There fore:  
Pilot project to look for possibilities for automatic  
(automated) coding**



# Source : Customs



Statistics Netherlands

- From the pre-arrival information of the Customs system SAGITTA the goods descriptions (free text format up to 255 chars) are selected.
- Indication: 1.000.000 different descriptions /year
- Numbers are growing because of
  - increasing number of ship-brokers that make use of the system
  - also data of outgoing container flow will be available

H150 CHEEZETONE

ROAD, NINGBO, CHINA

STUDLINKANCHOR 19X27.5M U1 STUDLINK ANCHOR CHAIN 40MM (E-EL-SLAC-EL-E) 100 X  
SPECIAL HDG RING 50X550MM

NITROSULFON- E

INDIAN NATURAL SESAME SEED

-192 CTNS 18,432 PCS OVAL HAT SHAPED TIN -62 CTNS 5,952 PCS OVAL HAT SHAPED TIN -89  
CTNS 6,408 PCS RECTANGULAR TRAY 255X192X35MM 1 DESIGN -341 CTNS 24,552 PCS  
RECTANGULAR TRAY

STATIONERY BATTERY T/S CARGO FM QINGDAO, CHINA TO ROTTERDAM VIA HONG KONG  
ENVIRONMENTALLY HAZARDOUS SUBSTANCE, SOLID, N.O.S.( LEAD)

20 PALLETS IQF FLAMED PEPPERS 10X10 1X10 KGS, 2.300 CARTONS

PART 2/2 109 ROLLS CONTAINING 11.910,50 MTS OF DENIM FABRIC TORIINVOICE: 16477 RE:  
07/0533601-001 SD: 2070402714/3FREIGHT PREPAID

PULLEY SPA

CNTR STC: 69 BXS METAL CUPBOARD, PADLOCK AND LOCK

CARBELAC 35 DOWNGRADED

RESINSOLUTION, NONHAZ 1 20' C/T CONTAINER CONTAINING: 7 PALLETS CONTAINING: 28 DRUMS RESIN SOLUTION, "NON HAZARDOUS"

STAPLE STAPLER STAPLE REMOVER

POY250D/48FSDRWAAGRADE POY 250D/48F SD RW AA GRADE

RIBBON CARTRIDGE

(SIXTY-FIVE) CARTONS OTHER POLISHES AND CREAMS 495.32K 1.047M

TANGERINS

CASES TEA, HERBS, PORCELAIN ITEMS & ADVERTISING MATERIALS

CNC PRESS BRAKE

MONGOLIAN NATIONAL GER

TABLELAMP TABLE LAMP

CIGARETTES

14 441 CARTONS CONTRACT NO : 60311300 2.900 CARTONS OF EACH 12/370 ML GLASSJAR YOUNG BABY CORN COUNT 13-20 PIECES OF REGULAR SIZE PER GLASSJAR PACKED IN 370 ML GLASSJAR WITH LITHOGRAPHIC MULTI-COLOURED SAFETY CAP NET WEIGHT 330G DRAINED WEIGHT 19...



# Basic principles



Statistics Netherlands

- Coding system general applicable (independent from production system and modality)
- Central processing
- Fine tuned with parameters
- Knowledge and rule based system.
- Growing and extendible system  
Not coding everything, but stop after a certain threshold
- Drawing a sample from remainder, codify it manually and gross up.



# Basic principles



Statistics Netherlands

- The aim is to codify automatically 60-75% of the goods automatically (representing 80-90 of the weight)
- Maximum Error: 1%
- Drawing a sample from remainder, codify it and gross up

**Automatic  
coded  
60-75%**

**Remainder**



# Method automatic coding



Statistics Netherlands

## Decision rules

- Using specific general available information (UN/HS)

## Algorithms

- Full text search
- Elimination of irrelevant information

## Databases

- Knowledge tables
  - Coding table (from description to code)
  - List of non relevant words
  - Signature strings





# Specific information : Signature strings



Statistics Netherlands

- Free text field sometimes contains widely used information :
  - Dangerous goods codes (UN-number)
  - Harmonised system (HS-code)
- Traceable codes using. SIGNATURE STRINGS, followed by a specific code

Dangerous goods:

***U.N.NO.:***

***UN#***

***UNNR***

***UN NUMBER:***

etc.

Harmonised System (customs)

***HS CODE***

***HS. NO.***

***H.S.CODE NO.***

***CUSTOMS COMMODITY CODE NO.***

etc.



# Dangerous goods



Statistics Netherlands

Description	Reference (NL)	NSTR	UNNR
(S.T.C.2 PALLETS) P-CHLOROPHENOL (UN#2020 CLASS:6.1 PACKING GROUP:III)	Chloorfenolen, vast	8190	2020
(HEXAMETHYLDISILAZANE) *IMO 3 UNNR 2924*	Vloeistof, brandbaar, bijtend, n.e.g.	8190	2924
(4,480 BAGS) RESORCINOL (1.3-DIHYDROXYBENZENE) SCE PO NO.4500000838 INV NO.ERP7568 "SEAWAY BILL" IMO CLASS:6.1 UN NO.:2876 PACKING GROUP:III LABEL:TOXIC(N) FLASH POINT : 164C TOXICITY : 980 MG (LD/TLV) EMERGENCY CONTACT	Resorcinol	8190	2876



# Harmonised system (HS/CN)



Statistics Netherlands

Description	NSTR	HS/CN
<b>(PALLETS) STEEL BOLTS WITH NUTS HS CODE 73181600 C Y TO CY FREIGHT PREPAID</b>	<b>9490</b>	<b>73181600</b>
<b>"MOON STAR" BRAND TITANIUM DIOXIDE NO.500R (20 PALLETS) H.S.CODE:282300</b>	<b>8190</b>	<b>282300</b>
<b>HI-LINE BEDFRAME SPF KD, HT, PET AS PER COMMERCIAL INVOICE GA0015 DATED OCT. 2/06 : HS COMMODITY CODE 4407.10.31 CLEARING AGENT: THE MAERSK CO. LTD MAERSK HOUSE, CLAN WILLIMTERRACE GRAND CANAL QUAY, DUBLIN 2 IRELANDCAED01A998VC198920061000768 SC NO: 92</b>	<b>0560</b>	<b>44071031</b>



# Relevant words: specific “popular” commodities



Statistics Netherlands

PRIO	string1	string2	Replace with
1	FROZEN	CHICKEN	
1	FROZEN	ORANGE PULP	
1	FROZEN	SALMON	
1	FROZEN	BEEF	
1	GARMENTS		
1	I.Q.F. STRAWBERRIES		FROZEN STRAWBERRIES
1	PATE DE CACAO		
1	POTATO	STARCH	
1	SEED	POTATO	
1	WHEAT	STARCH	
1	FR	PORK	FROZEN PORK
2	POTATOES		
2	SOY	BEANS	SOYABEANS
2	ORANGES		
2	PLYWOOD		
9	MAKITA		
9	NISSAN	MURANO	
9	VOLVO		



# Specific descriptions



Statistics Netherlands

Commodity description	normalised	reference (NL)	NSTR	HS/CN
- 16 CTNS = 319 PCS OF LADIES <b>JACKET</b> PO.06.06.846 STYLE.63510 MATERIAL 100% NYLON TASLAN WP/WR - 117 CTNS = 2,406 PCS OF LADIES JACKET PO.06.06.847 STYLE.69650 MATERIAL 100% NYLON TASLAN TD MC - 49 CTNS = 950 PCS OF LADIES JACKET PO.06.06.861 STYLE.612	GARMENTS	Kleding	9630	
"DAIKIN" <b>AIR CONDITIONER</b> ----- ***OOSTENDE TERMINAL 431 PACKAGES (417 CARTONS & 1 PALLE T & 13 CRATES)	AIR CON- DITIONERS	Airconditioning installatie	9390	84145990
1610 CARTONS ZIMBABWEAN <b>ORANGES</b> ON PALLETS	ORANGES	Sinaasappelen	0311	08051080
"YOKOHAMA" BRAND <b>TIRES</b> . K8502 200 PCS. K8636 600 PCS. K8565 200 PCS.. SC.NO TYO36000 . INTENDED MOTHER VESSEL:MAERSK SHEERNESS 0604 . (*) SWITZERLAND DOOR .	TYRES	Banden	9710	40122090
<b>ORANGES</b> - VALENCIA .	ORANGES	Sinaasappelen	0311	08051080
_ DETAILS AS PER ATTACHED SHEET - ** FAX:31-10- 4130029ATTN:MR.M.METHORST (INNER PKG: 2880 BL BALES) *SYNTHETIC RUBBER *(SOLUTION-POLYMERIZED STYRENE* <b>BUTADIENE RUBBER</b> )	BUTADIENE RUBBER	Rubber, synthetische butadieen-	0920	40022000



# Normalisation and codification



Statistics Netherlands

- Deleting of irrelevant words and elements..
- Linking normalised description with codification table (most important table of the system)

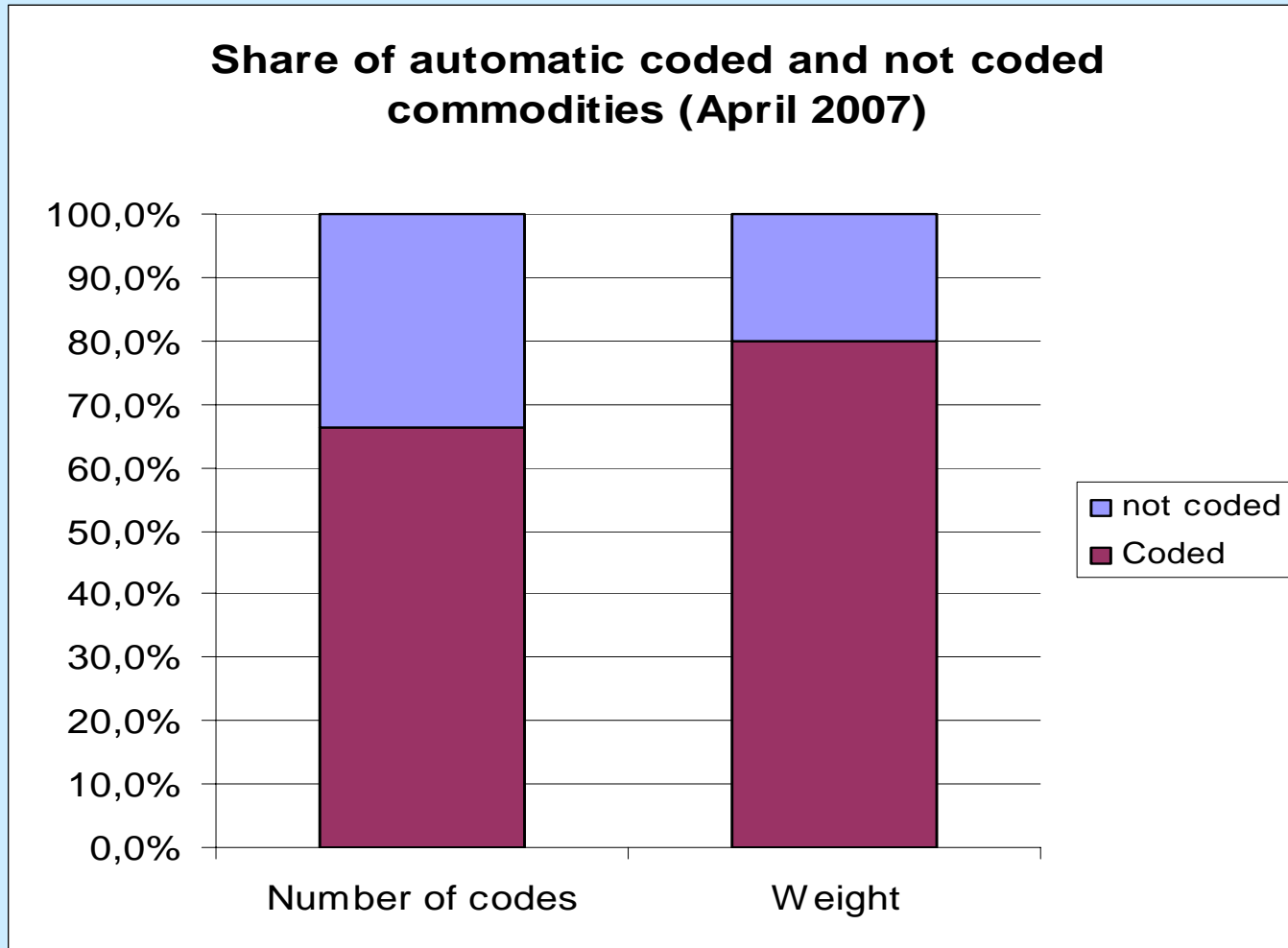
normalised	Reference (NL)	NSTR	CN/HS	UNNR	ADR	NST 2007
ACETIC ACID	Azijnzuur	8190				08.2
ACRYLONITRILE	Acrylonitril	8190	29261000	1093	30	08.2
ADIPINSAURE	Zuren, adipine-	8190	29171210			08.2
BARIUM TITANATE	Titanaten van onedele metalen	8190	28419090			08.1
CASHEW NUT KERNELS	Cashewnoten zonder dop	0359	08013200			01.4
CHRISTMAS PRODUCTS	Versieringsartikelen voor kerstmis	9790	95051090			13.2



# Results automatic coding



Statistics Netherlands



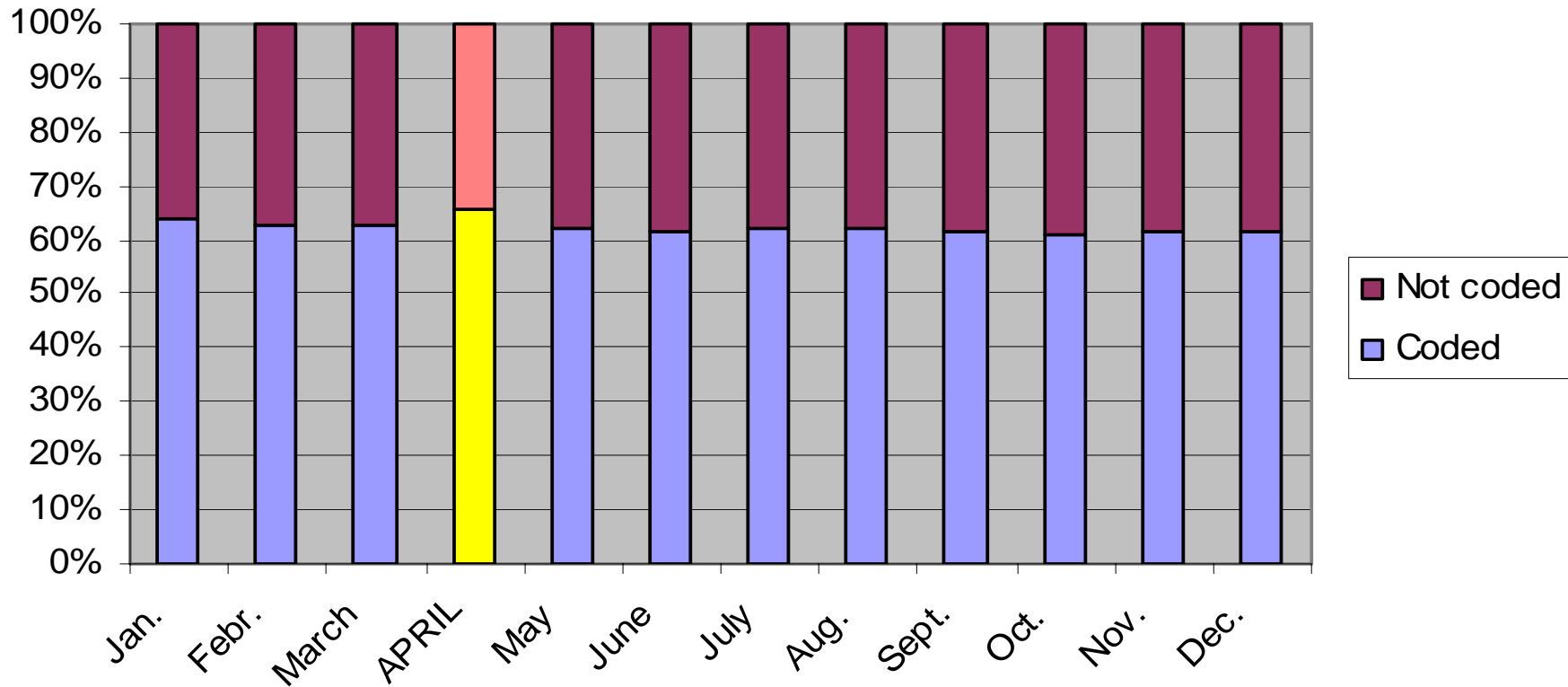


# Method applied on other months (1)



Statistics Netherlands

## Percentage of automatic coded goods descriptions in containers Jan-Dec 2007





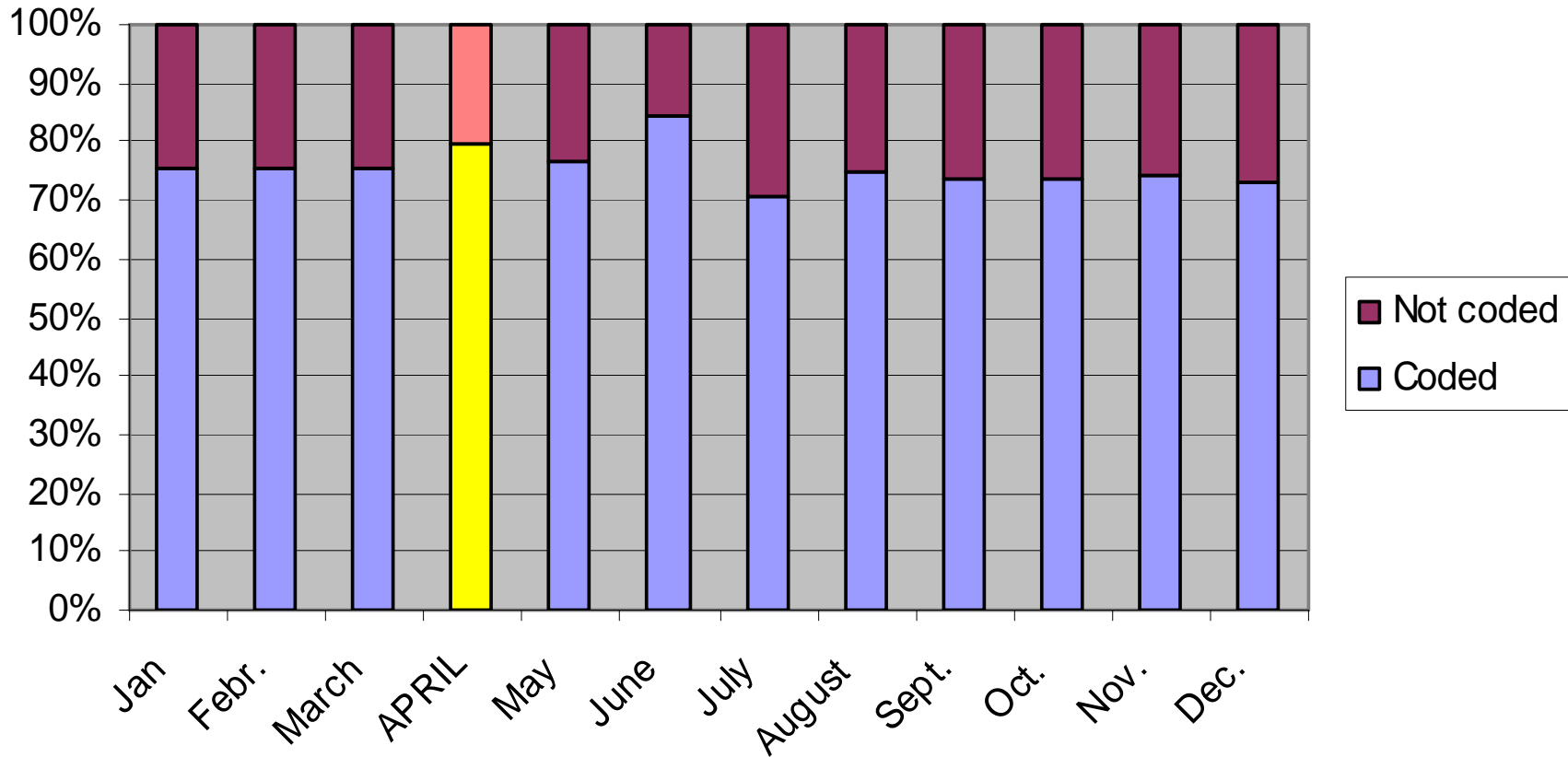


# Method applied on other months (2)



Statistics Netherlands

### Percentage of automatic coded weight of containerised goods jan-dec 2007





# Remaining (not coded)



Statistics Netherlands

- Sample of 10% from the not coded goods (about 3300 elements)
- Manually coding
- Imputation using the “random hot deck” method

**Automatic  
coded  
60-75%**

**Remainder**

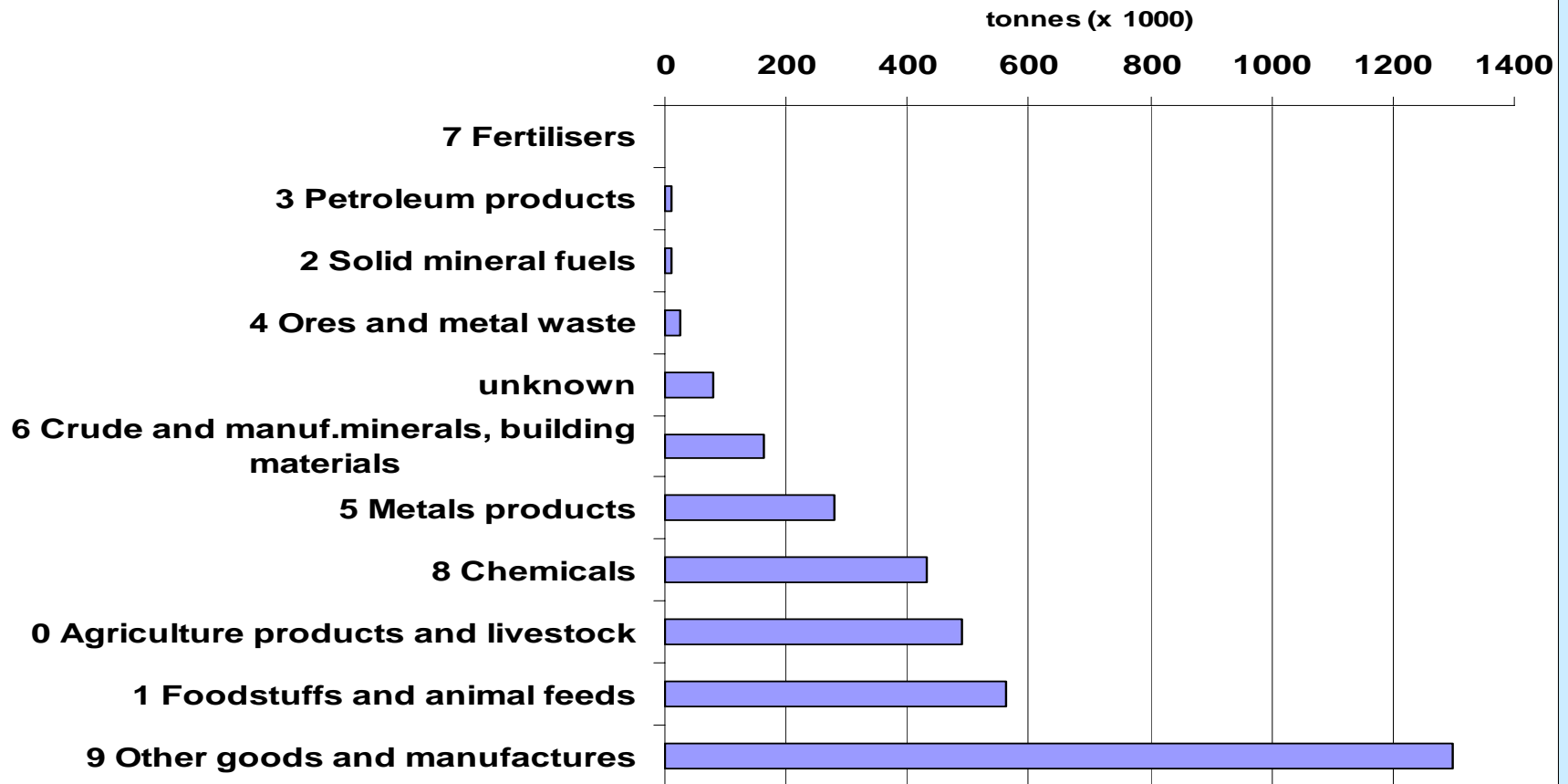


# Results



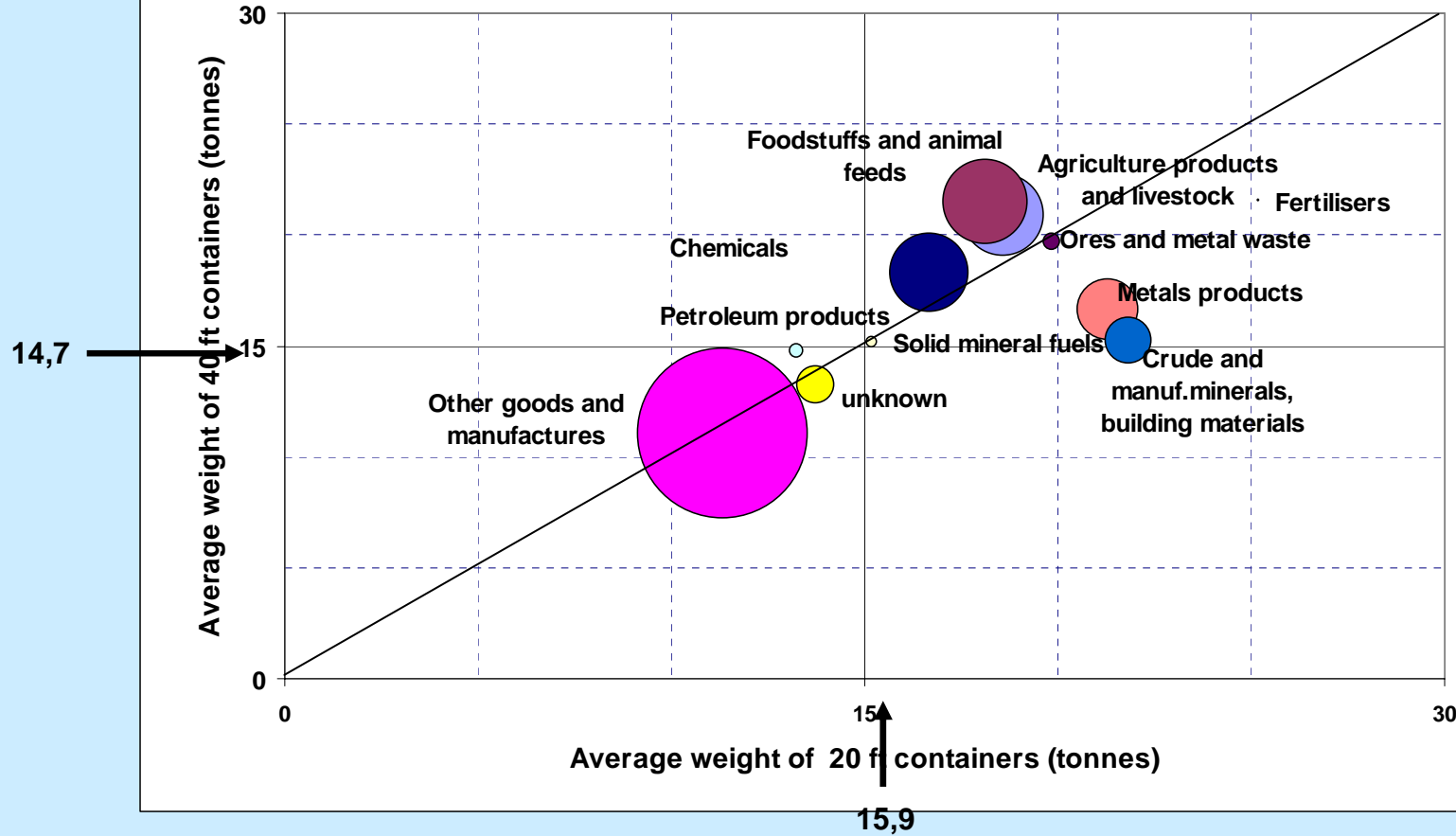
Statistics Netherlands

## Goods in containers by chapter NSTR, April 2007

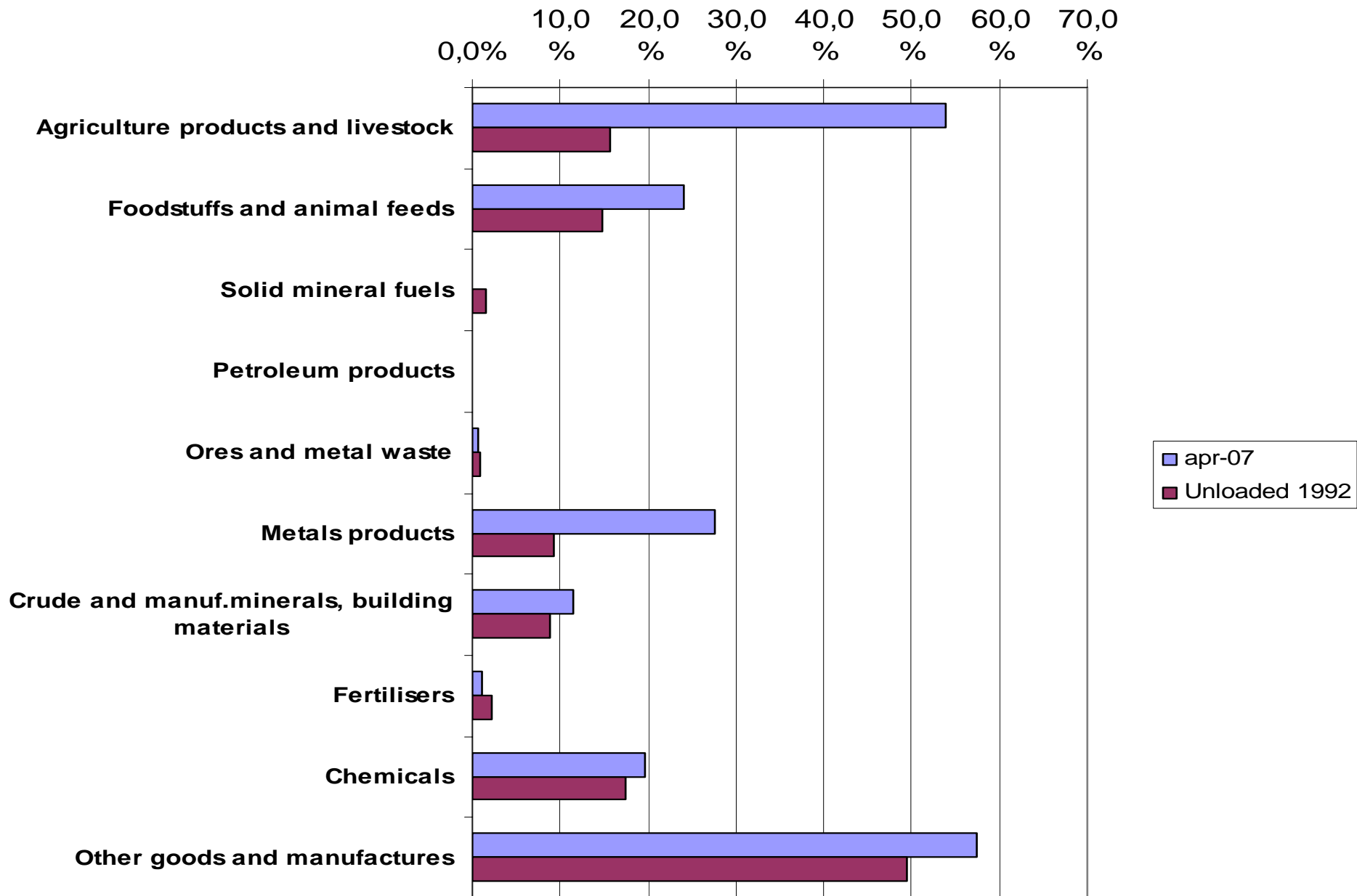




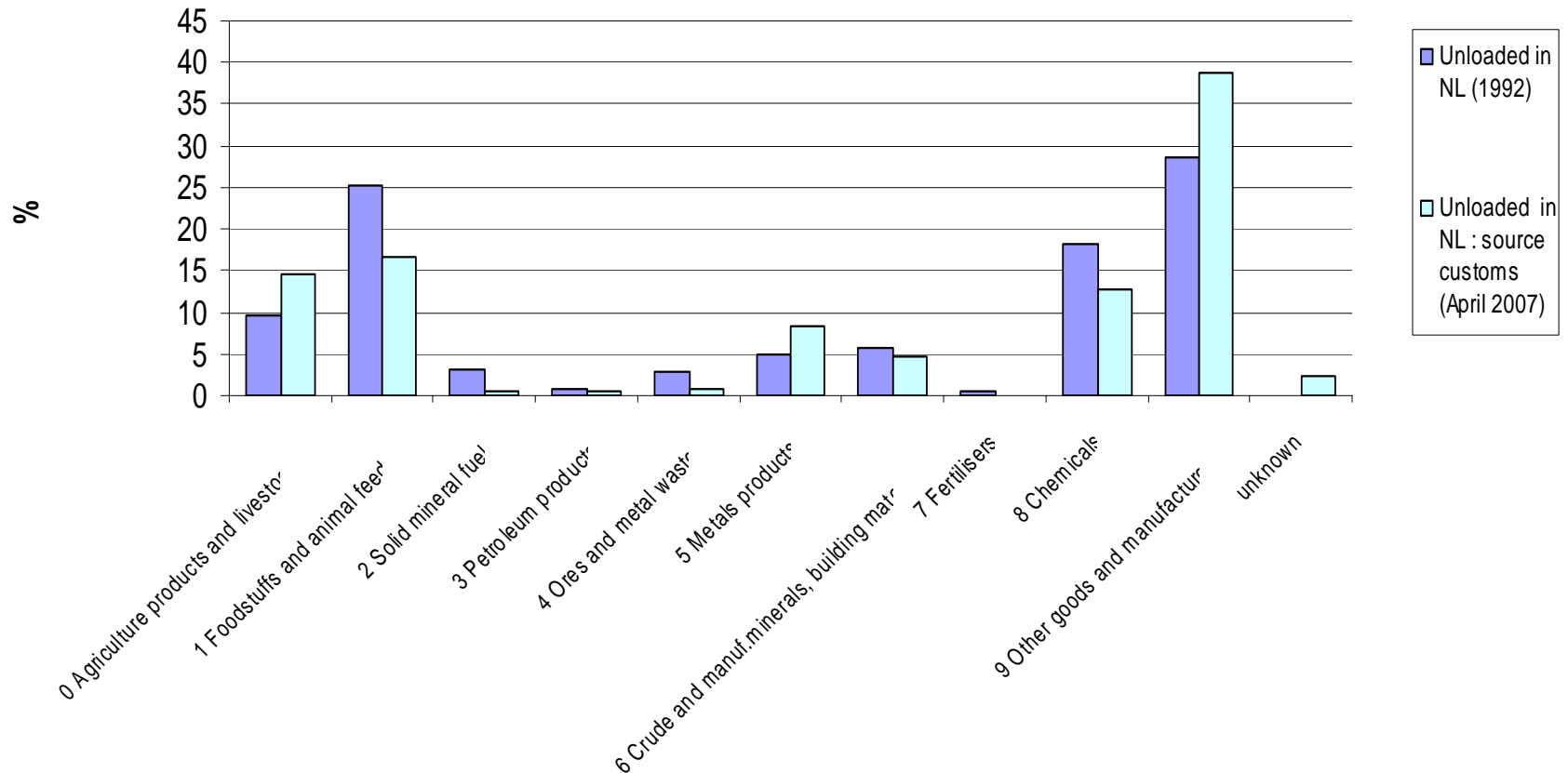
### Average weight of 20 and 40 ft containers based on number of containers per chapter NSTR



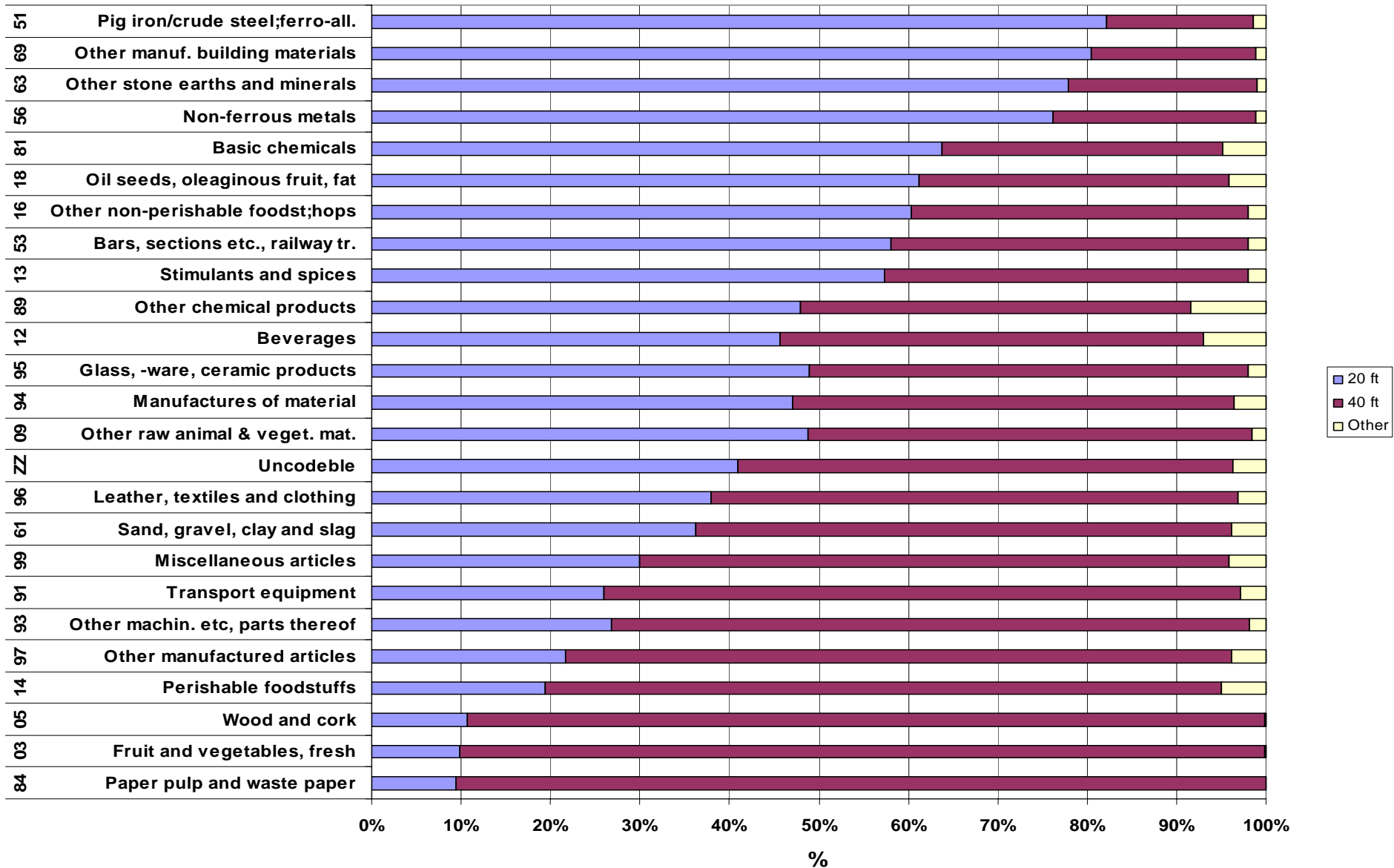
## Degree of containerisation, 1992 - April 2007, unloaded goods by chapter NSTR



### Container transport by sea Transit statistics 1992 and april 2007 compared



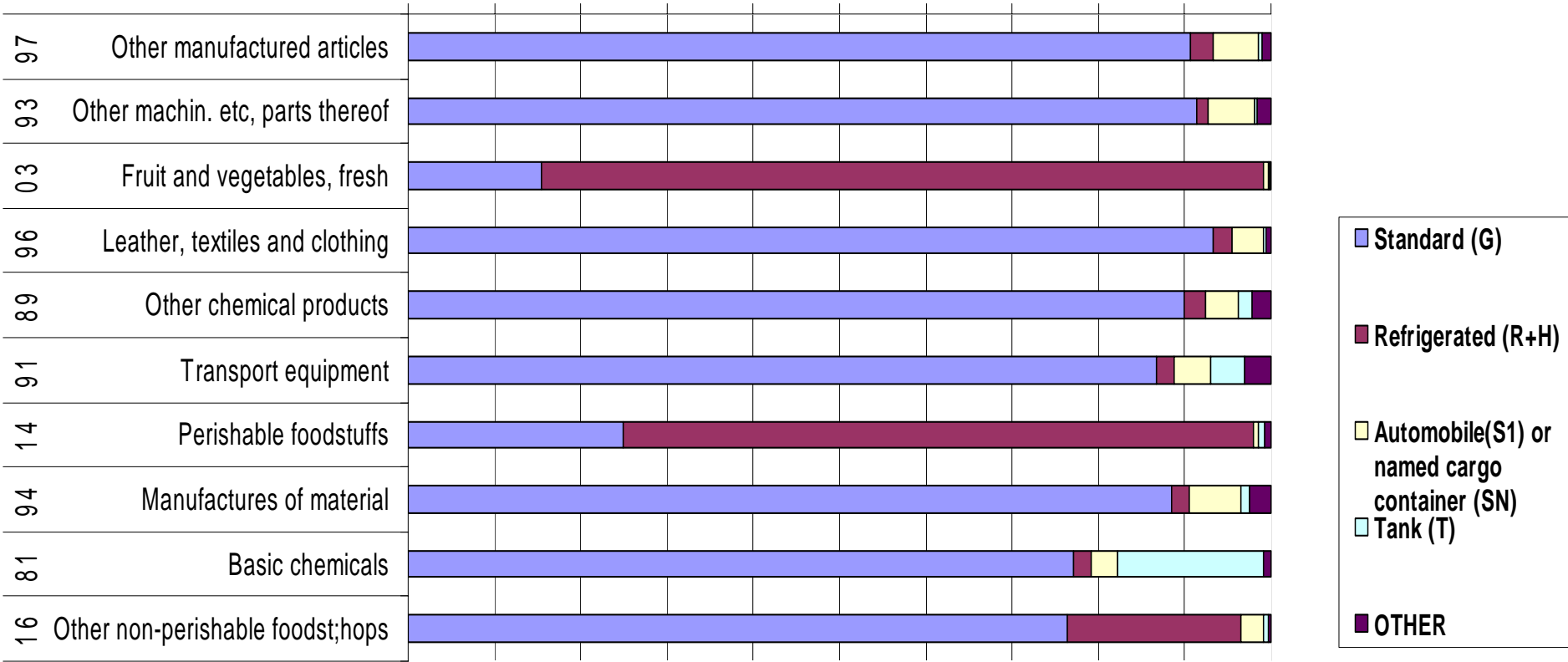
**TOP 25 commodity groups (NST/R) (based on number of containers).  
Relative share per container size class, April 2007**





### Top 10 NST/R Commodity groups, share of containertypes

0% 10% 20% 30% 40% 50% 60% 70% 80% 90% 100%





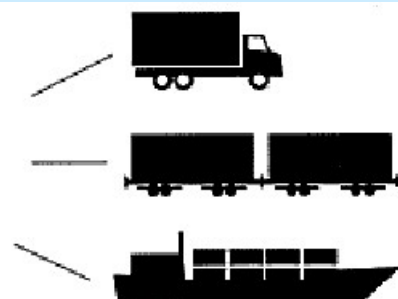


# Future developments



Statistics Netherlands

1. Decide on options:
  - Closing pilot and have data for just one month (April 2007)
  - Extending to one year and repeat it in 3 or 4 years
  - Incorporation in production process of maritime statistics
2. Cost benefit analysis of the options
3. Coding tables must be extended more  
This is very time consuming (expensive).
4. Improving sampling method
5. Enrichment of existing uni-modal statistics with commodity details (NSTR/NST2007) (Maritime, IWW, Rail)





- Thank you for your attention
- Questions or remarks?

**Peter Smeets**

**Statistics Netherlands**

**Division of Business Statistics**

**Statistical Analysis Department**

**Unit Traffic and Transport**

**P.O. Box 4481**

**6401 CZ Heerlen, The Netherlands**

**Phone +31 45 570 6432**

**E-mail: [psts@cbs.nl](mailto:psts@cbs.nl)**

