Swiss Federal Institute of Technology Zurich
SusTec
Sustainability & Technology

Universiteit Utrecht
Copernicus institute of Sustainable Development

Master's thesis

Projecting the uncertain future of mobility; Improving understanding of global transportation modeling

Submitted on:
Friday, July 19, 2013
by

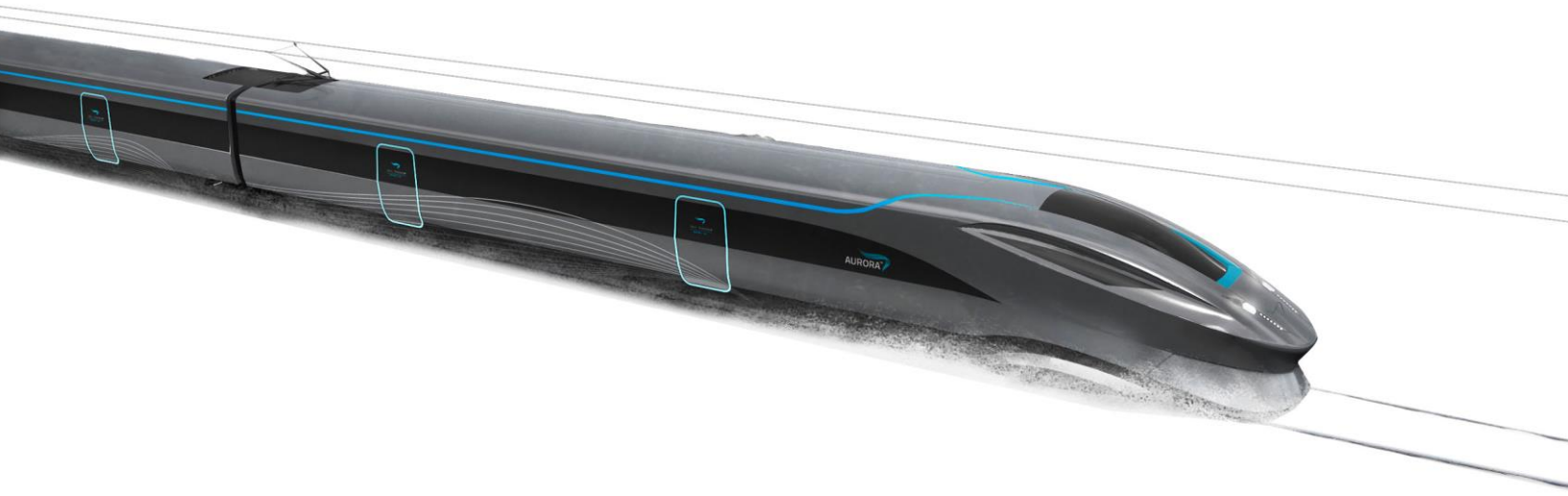
Roel Stijl,

MSc Candidate Energy Science, Utrecht University

Under supervision of:

Bert de Vries (Copernicus Institute, Utrecht University)
Bastien Girod (SusTec, ETH Zurich),

http://globaltravelmodels.codeplex.com



Intentionally left blank

EXECUTIVE SUMMARY

Citizens in continents from Africa to Asia to Europe have one thing in common, they spend between 1 and 1.5 hours per day, 5% of their lives, traveling. But how far and by what means do they choose to travel? Transportation is the third largest sector in terms of energy use. The last 30 years travel demand, the number of person km traveled per year has been growing at staggering 3.7% per year. With petroleum products, such as gasoline, diesel or kerosene, being virtually the only source of energy. With current technology travel demand therefore determines the CO_2 emissions caused by transport. Biofuels were hailed as a solution only a decade ago, but is not a silver bullet for the whole system, with land use issues and questionable CO_2 neutrality. Hydrogen cars and electric vehicles might be a new solution that could fill this gap. What is the impact of these innovations and economic development on travel demand?

Transportation modeling attempts to answer aforementioned questions. To this end many different types of models, with different approaches, levels of detail and predictions have been proposed. They differ in scope, the majority of models are designed for local policy forecasts and some attempt to forecast the future of the entire planet. No matter the model, if they are to be interpreted correctly it is crucial to understand how they function, how accurate they are, and how they relate to other models. Recently, several of these comparisons have been published, but by comparing the results produced by multiple authors the level of detail at which the models can be compared. This investigation recreates the travel modules of three established models, TIMER, GCAM and POLES. Moreover it creates a model based on commonly used econometric models in the field, the SIMPLE model.

Firstly the four models were distilled from publications, technical descriptions and cooperation with authors. The models were created in one framework with unified formulations, calibration methods, and analysis methods. Data is taken from datasets by Schafer, the TIMER model, open access database and literature research. The models were calibrated and validated on historic data, author published results, and empirical knowledge. A methodology was devised to determine ranges for Monte Carlo simulations in an unbiased manner. The Monte Carlo simulations ran on the Brutus supercomputer, #10 in EU, at ETH. The outcomes were used to produce probabilistic projections of travel demand, and implications for CO_2 emissions.

Secondly a comparison of the different modeling approaches is made. Results show that if the constraints of travel money budget and travel time budget are violated, then so are historic fits. Additionally there is evidence that competition based approaches perform better than per mode growth approaches such as elasticities per mode. It is also concluded that elasticities are more complicated to model than what is presented in the models here investigated; on this long a timescale they are time dependent. Time trends and saturation levels, applied to these elasticities don't remedy this. It is next found that a feedback mechanism for income on price is essential for a properly fitting model. Lastly the TIMER-travel and GCAM models are modified to converge consumer preferences of the developing region to the industrialized region values.

Thirdly projections are created, to this end the models were tested on historic validity. It was found that most model and region combinations fitted well historically and can be used to forecast the future. Some models, however, produced large errors compared to historic data for some regions. China was discarded all models and all developing regions for half the models. Next the projections were created, these were used to compare the forecasts of the different models and include the range in their inputs. Some of the results were that Industrialized region are likely to double TD in 2100 compared to 2005 and developing region are to increase this by at least 8-fold. The TD of latter will be some six times greater by 2100 than industrialized nations, compared to nearly equal TD for both today. Aircraft demand is to increase to least 12-fold.

Finally the investigation makes recommendations for future improvements on models, model comparisons and datasets. One of the main conclusions is that combining lessons learned could yield better models, perhaps by making a hybrid model. Also calibrations & model development should be done worldwide to include developing region. Additionally, improving datasets could yield more consistent results worldwide, this could be achieved by using local datasets.

Intentionally left blank

1 INHOUD

TABLE OF FIGURES

TABLE OF TABLES

TABLE OF ABBREVIATIONS

Abbreviations

Afr	Sub-Saharan Africa
CPA	Centrally Planned Asia
Dev	Developing region
EEU	Eastern Europe
FSU	Former Soviet Union
GCAM	GCAM Model
Ind	Industrialized region
LAM	Latin America
MC	Monte Carlo
MEA	Middle East & North Africa
MNL	Multi Nominal Logit
NAM	North America
PAO	Pacific OECD
PAS	Other Pacific Asia
POLES	POLES Model
Ref	Reforming region
SAS	South Asia
SIMPLE	Simplified elasticities model
TIMER-travel	Travel module of the TIMER model
WEU	Western Europe

Variables		Units
μ	Cost Distribution Parameter	-
COT	Cost of Time	$/pkm
ψ	Comfort Elasticity	-
CF	Correction Factor TMB	% of income
CO_2	Carbon Dioxide	-
DDS	Door to Door Speed	Km/hr
EC	Energy Costs	$\$_{2005}$/vkm
EE	Energy Efficiency	MJ/vkm

EP	Energy price	$\$_{2005}$/pkm
EQ	Correction for vkm of motor cycles	-
ER	Equipment Rate	vehicles / capita
ERSAT	Equipment rate saturation level	vehicles / capita
EREY	Elasticity on Equipment Rate	-
α	Elasticity on motor cycle TD	-
ϕ	Elasticity on motor cycle TD	-
σ	Travel Price Elasticity, (Total Cost of Ownership Elasticity)	-
δ	Price Elasticity, Fuel Price	-
$\delta 0$	Starting value for γ, in 1971	-
ϵ	Income Elasticity	-
$\epsilon 0$	Starting value for ϵ, in 1971	-
FP	Fuel Price	$\$_{2005}$/pkm
HW	Hours Worked	Hr/yr
I	Income, GDP/capita on PPP basis	$\$_{2005}$/cap, on PPP basis
α	Inertia	-
LF	Load factor	People / vehicle
m	Mode	-
MC	Mode Cost	$\$_{2005}$/pkm
MCR	Mode Cost Relative	-
MS	Mode Split	-
NEC	Non Energy Costs	$\$_{2005}$/vkm
NRSME	Normalized Root-Mean Squared Error	-
OC	Other Costs	$\$_{2005}$/vkm
PF	Preference Factor	-
r	Region	-
SW	Share Weight	-
t	Time	-
TCO	Total Cost of Ownership	$\$_{2005}$/vehicle
TD	TD	Pkm/capita/year
TDfit	Error on historic fit of TD	-
TDSAT	Saturation level of TD for HSM	Pkm/capita/year
TMB	Time Money Budget	% of income
TMBA	TMB added factor	% of income
TP	Travel Price	$\$_{2005}$/pkm
TTB	Time Travel Budget	Hr/year
TTC	Total Travel Cost	% of income

TU	Time Use	Hr/year
TW	Time Weight	-
VTD	Vehicle TD	Vkm/yr
WR	Wage Rate	$\$_{2005}$/hr
λ	Substitution elasticity	-
s	Sigma	-

1 INTRODUCTION

In academic circles it is now widely accepted that anthropogenic greenhouse gas (GHG) emissions are affecting the planet at an unprecedented scale (IPCC, 2007). The main source for these emissions is directly linked to the usage of energy (EAR, 2012). During the last decades initiatives have spawned that attempt to reduce energy consumption and emissions, such as new building regulations, emission trading and targets for biofuels (International Energy Agency, 2012). For these types of initiatives to work successfully, a thorough understanding of these systems is necessary. Therefore the topic of this investigation is to improve understanding of one of these systems, the global transportation system, a system that is growing at 3.7% annually for the last 30 years and is therefore of growing concern (Zhang, Jiang, & Liu, 2007).

This section aims to first introduce the energy and emission challenges. Then it will zoom in to the specific field of passenger transportation modeling. Lastly it will introduce the research question and scope of the present investigation.

1.1 ENERGY AND EMISSION CHALLENGES IN TRANSPORTATION

The transportation sector is a significant consumer of energy (EAR, 2012), using up to 24% of primary energy in the USA in 2012 (EAR, 2012) and is hence a target for reduction in energy use as well as GHG emissions. Unlike the worldwide focus of this investigation, in the past the focus of the transportation problem was thoroughly placed in local pollution. Such as issues with smog, fine particles (Kupiainen & Klimont, 2007) and even the great London manure problem (Davies, 2004). These issues remain in many metropolitan areas, however there are areas such as in Switzerland, where the pollution is less of a problem. It has reached such low levels that further policy no longer improves the situation in an efficient way (De Haan, 2009). In this light solutions to the local pollution problems hence exist (Kupiainen & Klimont, 2007).

The reduction in these local pollutants can to a large extent be attributed to the successful implementation of policy aimed at the manufacturers. This led to improvements in technology that enabled the cleaner burning of fossil fuels. GHG emissions, however, are inherent to the use of fossil fuels. Although the efficiency of engines of all types has gone up considerably since their first incarnation (Schafer, Heywood, Jacoby, 2010), their widespread adoption in the last century now makes transportation the number three sector in terms of energy use and GHG emissions in the world (International Energy Agency, 2012). Reducing GHG emissions globally will therefore require novel approaches quite different from the instruments used to reduce local pollution in the past.

Historically road transportation has developed from human powered transportation modes to animal powered transportation toward motorized transportation (Schafer, Heywood, Jacoby, 2010). The past century saw the mass adoption of motorized vehicles, including light duty vehicles (LDV) and motorized public transportation (PT) in the industrialized regions. With increasing wealth motorized transportation is starting to pick up in the developing regions. First

walking or biking will be replaced by mostly public transportation and motorcycles (Kitous, 2010a), but at a threshold income of approximately 5000$_{2010}$/capita these modes will be replaced by a growing stock of cars (Chamon, M. Mauro, P. Okawa, 2008) and in the future likely planes or high speed train (Schafer, Heywood, Jacoby, 2010).

Current technologies such as electric vehicles allows ground-based transportation modes to be less CO_2 intensive, hypothetically even CO_2 neutral, and hence reduce future emissions (Deetman et al., 2013). This possibility is granted not just by an increase in the efficiency of vehicles with traditional combustion engines but rather by the ability to shift to more sustainably produced forms of fuel. For example electricity, hydrogen, or biomass (Edenhofer et al., 2011). Such a transition is not at all unfeasible as can be seen in the historic shift, be it with climate unrelated motives, in train fuels from coal to diesel to electric (Schafer, Heywood, Jacoby, 2010). A recent modeling study in fact shows that such a transition is possible without losing much of the mobility worldwide (Bastien Girod, Van Vuuren, & Deetman, 2012).

Ground-based transportation is one thing; another altogether is that of air transportation. In the decades after the Second World War the use of transportation has taken for the sky. With unprecedented door to door speeds and plummeting prices, airplane transportation is the fastest growing sector of transportation by a margin, averaging 9% growth per year in the USA compared to 2% growth for cars (Hepburn & Müller, 2010). Unlike other modes of motorized transportation, cars, busses, trains and even high-speed trains; at present low emission aircraft are not technologically feasible. Alternatives such as hydrogen- (Westenberger, 2008) or even nuclear (Khandelwal, 2011) aircraft have been proposed but are uncompetitive and at present technically unfeasible.

In addition the increase in travel demand for all modes of transportation in the Developing regions, is a growing issue (Wright & Fulton, 2005). These regions can, by sheer volume, negate all the improvements in emission standards the industrialized world puts out.

1.2 TRANSPORTATION MODELING

The topic of the present investigation, transportation modeling, is generally part of broader field of energy modeling. In order to investigate the future direction of the energy system aggregated energy models were made by a variety of authors, such as TIMER (De Vries, Van Vuuren, Den Elzen, & Janssen, 2001), GCAM (Edmonds, Wise, & MacCracken, 1994), POLES (A. Kitous, 2000) and GET (Grahn, 2007). These models rely on sub models that simulate each relevant facet of the energy system. One of these facets is the transportation system, generally split up in freight and passenger transportation.

The modeling of the transportation sector attempts to understand the past, present and future trends of passenger transportation in regions worldwide. More specifically how many kilometers people (pkm) and vehicles (vkm) will travel, the travel demand TD (pkm/cap/yr). Additionally it is investigated which modes, such as cars, aviation and train, are used. At a more detailed level the modes are disaggregated into different vehicles that use different fuel types at different efficiencies. With this information one can forecast CO_2 emissions and energy usage. These models can for example also be used for integrated assessment studies, (e.g. (Hepburn & Müller,

2010)), understanding traveler behavior (e.g. (Bastien Girod et al., 2012)) or forecasting future developments (International Energy Agency, 2007).

Worldwide travel behavior follows trends that depend on macro-economic variables such as national income, population density and the price of energy. Additionally several empirical variables are known to be stable at a very aggregate level (Zhang et al., 2007), these most well-known of these variables are the travel time budget[1] (TTB) and travel money budget (TMB). Denoting the tendency for people to spend a fixed amount of time in transit and the tendency for people to spend a fixed percentage of income on traveling, respectively. Both trends are displayed in Figure 1. The definition of TMB used in this publication is dependent on the share of cars, increasing from approximately 3.5% of GDP to approximately 12% of GDP[2] (Schäfer & Jacoby, 2005). The TTB, including non-motorized modes, is fixed, at 1.0 – 1.5 hours per capita per day (Susilo & Avineri, 2011), without motorized modes it increases with income from 0 to 1.0 – 1.5 hours per capita per day (Avineri, 2009). Given these constraints economic development tends to drive people toward faster modes, effectively optimizing the distance traveled within the budget (Schafer, Heywood, Jacoby, 2010). Some authors additionally add a comfort variable (Bastien Girod, Van Vuuren, & De Vries, 2013) or value of time (Kyle & Kim, 2011) to this spending paradigm.

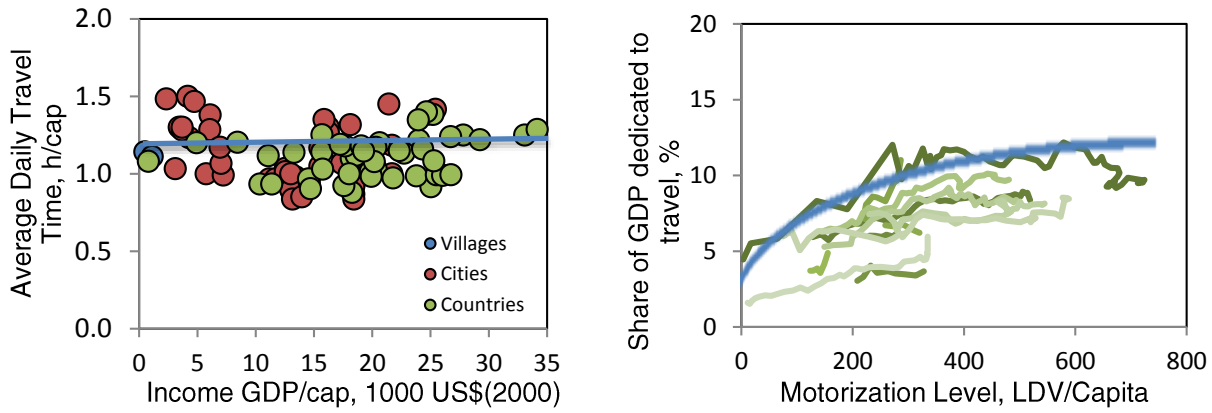


Figure 1: Travel money budget (right) and travel time budget for motorized and non-motorized modes (left) (Schafer, Heywood, Jacoby, 2010). The blue line indicates the values used.

The modeling work already done in this field is extensive. Suganthi categorizes the models in a convenient manner (Suganthi & Samuel, 2012). The models can be categorized into these model types:

- Regression based models (Schafer & Victor, 2000)
- Computable general equilibrium (CGE) models (Schäfer & Jacoby, 2005)
- Time series models (POLES) (A. Kitous, 2000)

[1] There is a debate whether or not TTB actually this fixed, without aggregation this is certainly true. For example cities in china are known to deviate strongly between each other, but at the aggregate level countries are quite similar. In order to avoid this debate the value is taken as a range.

[2] This is introduced based on empirical findings and might deviate especially for less developed countries. In this investigation a broad range is therefore assumed or a correction for regional differences is used.

- Econometric models (SIMPLE)
- Decomposition models (TIMER-travel) (Bastien Girod et al., 2012)
- A combination of these, a hybrid model (GCAM) (Kyle & Kim, 2011)

Additionally there are models that work at a local- (Litman, 2011), nation-wide- (Zhang et al., 2007) and at a global level (Grahn, 2007). Because these studies have implications for, for example, infrastructure projects and nation-wide policy there is a plethora of local and nation-wide investigations. Many of these looking into specific modes of transportations and their dependency on from example the environment, behavior and cross elasticities (Litman, 2012a).

The global travel models are scarcer and less diverse in their approaches. Based on a study into different global travel models (Bastien Girod, Van Vuuren, Grahn, Kitous, & Kyle, 2013) it is decided to analyze three models that use a decomposition approach (Bastien Girod et al., 2012), time series approach (Kitous, 2010a) and a hybrid approach (Kyle & Kim, 2011). These models are derived from the energy wide forecast models TIMER, POLES and GCAM, respectively, see Table 1. Additionally a model is created to represent popular econometric methods used in many publications in the field, the model is called SIMPLE.

Hence from the transportation modules of these models the TIMER-travel model, the POLES model, and the GCAM model are constructed. Additionally a simple elasticity model (SIMPLE) is constructed based on the extensive literature research into elasticities, largely based on the work of Litman (Litman, 2011).

Characteristic	TIMER-travel	GCAM	POLES	SIMPLE
Mode competition	Yes – MNL type	Yes – Logit type	Bus & car on (gasoline)fuel price	No
Summated TD	TMB * Travel price	Elasticity on summated TD	Elasticity per mode	Elasticity per mode
TD saturation	Limited by TMB	None	Saturation on LDV and HSM	None
Travel price used in TD	Travel price + comfort variable	Travel price + cost of time	Fuel price	Travel price
Travel price used in mode split	Travel price + Preference factor & time weight	Travel price + cost of time	Modes are independent	Modes are independent

Table 1: Key characteristics of different models

1.3 RESEARCH QUESTION AND SCOPE

The present investigation focuses on investigating the differences between different approaches in modeling the global transportation system using a single modeling framework. Three well-established models, GCAM, POLES and TIMER-travel, and a model recreating common econometric methods, SIMPLE, are compared. The primary method of investigation uses probabilistic projections and statistical methods to analyze each model under a given range in input parameters. By performing the analysis in the same modeling framework, using the same input parameters, datasets and calibration procedures, the models are compared in a manner that is kept invariant of the model (Botterweg, 1995). This and the probabilistic projections are the novelty of the approach.

The research question of the present investigation is hence, "How do different modeling approaches compare in representing past global travel patterns and what are the implications in the future?".

This can be divided up into the following sub-questions:

1) What modeling approaches are suitable for comparison and how can they be implemented?
2) How do the models compare in replicating historically known facts such as the actual TD and time travel budget?
3) What can the use of probabilistic projections tell us about the development of TD, emissions and energy use in the coming century?

The research scope is:

- To replicate several models with a significantly different modeling approach
- To analyze the individual models and compare their strength and weaknesses in representing history
- To compare the different future projections by using probabilistic projections and statistical methods
- To come up with suggestions for enhancements to existing models and if possible create enhancements

The methodology and structure of the present thesis is based on the three questions posed above. In short form and in order: How do the models compare conceptually? How do the models differ on historic replication? How do the models compare in terms of future projections?

2 METHODS

The basic methodology of this investigation is based on the general process of model development(Katz, 2002), calibration, validation and analysis found in many modeling approaches including TIMER (De Vries et al., 2001) and IMAGE (Bouwman, Kram, & Goldewijk, 2006). Because multiple models are compared, time is spent building a general modeling framework and unified procedures for all models. The overarching methodology of the present investigation is displayed in Figure 2, the research questions are integrated into the method and the steps match the paragraphs in this section.

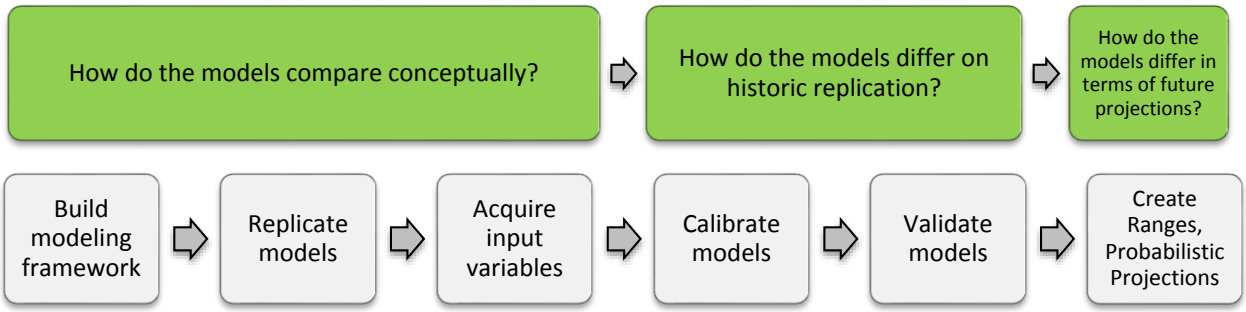


Figure 2: Flow chart of the general methods of this investigation, the top flow chart denotes the research questions, the lower one denotes the methods used to investigate them, the steps of the investigation match the research questions as depicted

The first step of this investigation is to build a general modeling framework in Matlab (Mathworks, 2012) that allows analysis of the models under identical conditions, such as datasets, calibration procedures and post processing tools. The framework is essentially the analysis tool custom-made for this investigation. The second step is to reconstruct each model based on literature provided by the authors. The third step is then to acquire and construct the required datasets. The fourth step is to validate the models based on historical data, subject to sensitivity analyses and compared against published results. The fifth step calibrates the model. The sixth step is to determine valid ranges for the probabilistic projections using the values from the calibration, a literature investigation, values from the authors of the model and Monte Carlo methods (Ruijven et al., 2009). Lastly these ranges are used to create probabilistic projections.

The probabilistic projections, combined with other analyses generated during the process, are used in order to compare each of the models used in the present investigation.

2.1 GENERAL MODELING FRAMEWORK

In order to compare each of the models before in a transparent manner a general modeling framework is constructed in Matlab (Mathworks, 2012). The framework at a low level of detail is described in Figure 3.

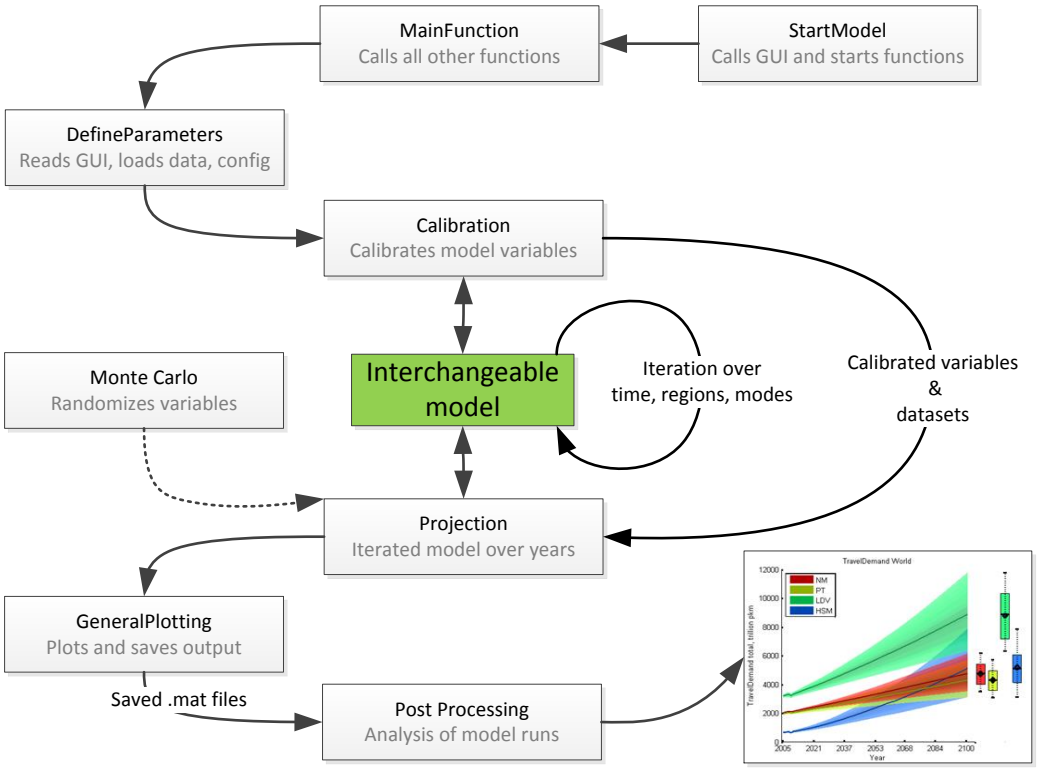


Figure 3: Layout of the Matlab modeling framework

The variables in the model framework are split up according to type. This is in order to facilitate sensitivity analysis, probabilistic projections and make the thesis and code easier to understand. The four classes of variables are the dependent-, independent-, exogenous- and configuration parameters. This notation will be used throughout this thesis.

Examples of each type of variable, definitions of the exampled variables can be found in the next section:

- **Dependent variables** are variables that are changed through interaction with the model, such as travel demand and errors.
- **Independent variables** are fixed before the model starts and include calibration parameters and constants. Examples are the elasticities and share weights.
- **Exogenous variables** are derived from datasets that are imported. Examples are the income datasets.

- **Configuration variables** contain settings of the models. These include, for example, the logical switches for settings and the definition of regions. They are only useful for Matlab internally

The interface of the framework, see Figure 4, is designed with a GUI in order to operate the model in a quick, user friendly and less error-prone manner. Critical parameters clearly visible and changeable. Additionally live tracking of modeling progress, optimization and graphical output enables quick debugging and improves model understanding. The interface of the post-processing toolbox is presented in Figure 5, each of the analysis can be performed from this interface and exports to formats and sizes required for publications are easily accessible.

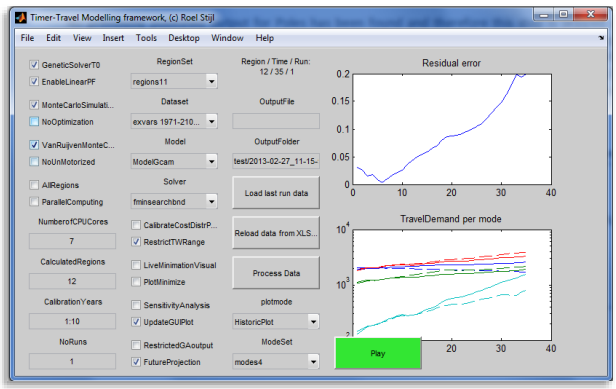


Figure 4: Interface of the modeling framework

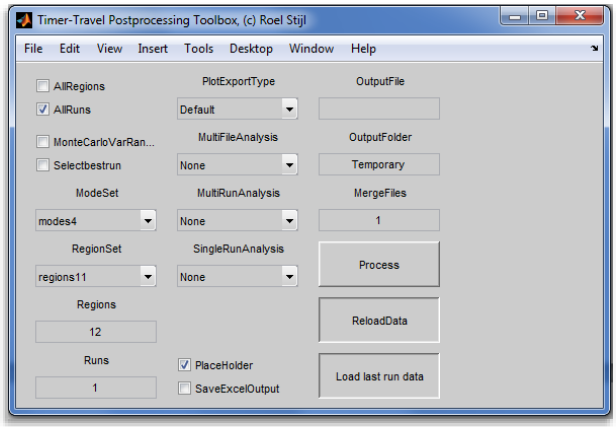


Figure 5: Interface of the post-processing toolbox

The models are constructed in a way that leaves them fully interchangeable. They are iterated over regions (r), time (t) and modes (m). The remaining dimensions, vehicle type and fuel type assume scenarios from TIMER (Bastien Girod et al., 2012) and are recalculated in the vehicles module, see section 2.1.1.

More information on the modeling framework can be found in the appendix, and additionally it is available open source from http://globaltravelmodels.codeplex.com, under the MIT open source license.

2.1.1 REGION CONVERSION

The published regions and modes used differ substantially between the models and the dataset. Therefore a conversion to a single set of regions is necessary for comparison of the models. This section will explain what conversions are performed. For the modes it is decided to use the 4 modes defined by Schafer (Schafer, Heywood, Jacoby, 2010). For the regions it is decided to use the 11 regions originally defined by Schafer (Schafer, Heywood, Jacoby, 2010), as shown in Figure 6.

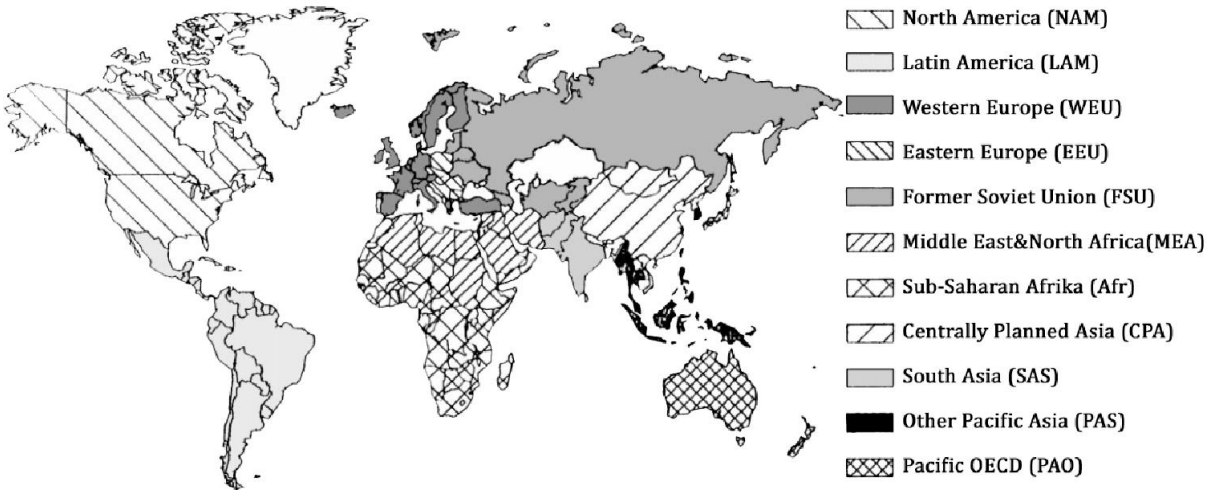


Figure 6: 11 regions used in the present investigation based on (Schafer, Heywood, Jacoby, 2010)

The original TIMER-travel model uses 26 regions, GCAM uses 14 regions, POLES uses individual regions, SIMPLE is defined irrespective of regions. The regions are compatible and conversions between each system can be made. The conversion used are depicted in Table 2 and Table 3. The aggregated regions used by Schafer and GCAM are depicted in Table 4.

Aggregation happens through linear matrix operations developed in the modeling framework (Louviere & Woodworth, 1983). Care had to be taken because future travel demand is unknown and hence some variables could not be divided into regions. Therefore some sets are reconstructed from the bottom up, discussed in the next section. Notice that the world region is used in two different contexts Firstly in use as a region it is the combination of all the input datasets, and used for calibration purposes. Secondly, in terms of an aggregated region, it is the outputs of the other regions added up. Differentiation comes from context.

Region in TIMER	Region in Schafer	Region in TIMER	Region in Schafer
Canada	North America (NAM)	Ukraine +	Former Soviet Union (FSU)
USA	North America (NAM)	Asia-Stan	Former Soviet Union (FSU)
Mexico	Latin America (LAM)	Russia +	Former Soviet Union (FSU)
Rest Central America	Latin America (LAM)	Middle East	Middle East and North Africa (MEA)
Brazil	Latin America (LAM)	India	South Asia (SAS)
Rest South America	Latin America (LAM)	Korea	Other Pacific Asia (PAS)
Northern Africa	Middle East and North Africa (MEA)	China +	Centrally Planned Asia (CPA)
Western Africa	Sub-Saharan Africa (AFR)	South East Asia	Other Pacific Asia (PAS)
Eastern Africa	Sub-Saharan Africa (AFR)	Indonesia +	Other Pacific Asia (PAS)
South Africa	Sub-Saharan Africa (AFR)	Japan	Pacific OECD (PAO)
OECD Europe	Western Europe (WEU)	Oceania	Pacific OECD (PAO)
Eastern Europe	Eastern Europe (EEU)	Rest of S. Asia	South Asia (SAS)
Turkey	Western Europe (WEU)	Rest of S. Africa	Sub-Saharan Africa (AFR)

Table 2: Translation table between Schafer and TIMER regions

Region in GCAM	Region in Schafer	Region in GCAM	Region in Schafer
Africa	Sub-Saharan Africa (Afr)	Japan	Other Pacific Asia (PAS)
Australia & NZ	Pacific OECD (PAO)	Korea	Other Pacific Asia (PAS)
Canada	North America (NAM)	Latin America	Latin America (LAM)
China	Centrally Planned Asia (CPA)	Middle East	Middle East & North Africa (MEA)
Eastern Europe	Eastern Europe (EEU)	Southeast Asia	South Asia (SA)
Former Soviet Union	Former Soviet Union (FSU)	USA	North America (NAM)
India	South Asia (SA)	Western Europe	Western Europe (WEU)

Table 3: Translation table between GCAM and Schafer Regions

Aggregrated Region	Region in Schafer
Industrialized	North America, Pacific OECD, Western Europe
Reforming	Eastern Europe, Former Soviet Union
Developing	Centrally Planned Asia, Latin America, Middle East & North Africa, Other Pacific Asia, Sub-Saharan Africa
World	Centrally Planned Asia, Latin America, Middle East & North Africa, Other Pacific Asia, Sub-Saharan Africa, Eastern Europe, Former Soviet Union, North America, Pacific OECD, Western Europe

Table 4: Translation between Schafer regions and aggregated regions

2.1.2 MODE CONVERSION

The modes of transportation also differ per model. Again a conversion to a single set is used. SIMPLE, GCAM and TIMER-travel are not principally defined for any modes in particular, although the latter two indicate the modes they used in their publication. They are defined as a

set of equations that treat modes symmetrically and are therefore relatively invariant to the modes used. The POLES model on the other hand is defined asymmetrically and hence has a separate equation for each of the modes. Additionally this model first calculates total public transport and then disaggregates to rail and bus. The POLES model is therefore is the restricting variable in the definition of the mode set. Lastly, Schafer (Schafer & Victor, 2000) omits high speed train in favor of only using air, he names this mode "high speed modes". It is therefore decided to use the modes non-motorized, public transport, light-duty vehicles and high speed modes. As described in Table 5. Non-motorized isn't used in the comparisons for it is not represented in POLES.

Model	**NM**		**PT**		**LDV**	**HSM**	
This investigation	Non -Motorized		Public Transport		Light-Duty Vehicles	High Speed Modes	
TIMER-travel	Walk	Bike	Bus	Train	Car	HS Train	Air
GCAM	Not defined but used		Bus, Rail		LDV	Air, HS Rail	
POLES	Not present		Bus	Rail	Car & Motorcycle[3]	Air	
Schafer	Not present		Bus	Rail	Auto	High Speed Modes	

Table 5: Translation between modes for different models, Schafer refers to the Schafer dataset

[3] Motor cycles are omitted because no data is available on them in Schafer's set

2.2 MODEL REPLICATION

This section describes the design of each model individually. Each model is replicated as closely as possible to the originally published models, whilst retaining a general notation and using a universal dataset. The models are discussed in an order that will be used throughout this thesis. The logic is to start with the simplest model and work toward the more complicated models. The last section will describe the modifications made in order to enable a convergence of the consumer preferences in the developing region in the models to the first world standards.

1. SIMPLE — The SIMPLE model is a recreation of the conceptual model used in most econometric analyses in transportation research literature. It uses elasticities on the travel pricing and income for each mode to determine travel demand growth.
2. POLES — The POLES model (A. Kitous, 2000) is a comprehensive recreation of the global energy landscape. The POLES travel submodel is part of this. It uses the same basic formulation as SIMPLE but puts several modifiers on the travel demand growth that are different per mode.
3. GCAM — The GCAM model (Kyle & Kim, 2011) is a comprehensive recreation of the global energy landscape. GCAM travel is a submodel of this. GCAM assumes an elasticity similar to SIMPLE but over all modes. It also add a competition between modes, consumer preferences and income dependent pricing.
4. TIMER-travel — The TIMER model is a comprehensive recreation of the global energy landscape. (Bouwman et al., 2006; de Vries et al., 2001). The TIMER-travel model (Bastien Girod et al., 2012) is a submodel. It does not primarily use elasticities like the other models. It instead determines travel demand using travel time budget and travel money budget. It includes competition between modes, consumer preferences and income dependant pricing.

As mentioned, the models have several independent dimensions that are iterated over. They can be seen as either dependencies of the variables or, more accurately in the programming the dimensions of matrices. The dimensions are region (r), time (t), modes (m). The latter consist of non-motorized NM, public transport PT, light duty vehicles LDV and high speed modes HSM, a closer description can be found in section 0. The notations of the authors have been modified to reflect this notation, but left close to the original publications as possible for easy referencing.

The output of each model is the travel demand TD (pkm/cap/yr) per mode. The mode split MS is the fraction of TD assigned to each mode. TD and MS are related as follows:

$$MS_{r,t,m} = \frac{TD_{r,t,m}}{\sum_{modes} TD_{r,t,m}} \; [-] \qquad (1)$$

2.2.1 THE SIMPLE MODEL

SIMPLE will be considered as a reference for the other models. It is the simplest model conceivable in the field of transportation modelling. It is based on the most frequently published method for research by local municipalities, large infrastructure projects and other types of applied transportation research (Litman, 2012b). The popularity of this type of analysis is due to its simplistic derivation. A simple linear regression on a log-log projection of data is sufficient. The results are valid in the short run, several year and the concept is intuitively easy to understand.

The econometric models used in these analyses are the model presented. The growth in travel demand TD (pkm/cap/yr) is determined by elasticities applied to income and pricing. Notice that there is a distinction between short run- and long run elasticities. The difference is best illustrated with an example, if fuel prices increase 50% the immediate effect might be a small change in travel demand for LDV as people don't change behavioral patterns quickly. In the long run the same price increase will lead to a larger reduction in TD. Short run elasticities represent the effect over several years and long run elasticities represent the effect over longer periods of time. The latter is therefore used in all references to elasticities.

At the core of this model is a simple relationship of elasticities. The income elasticity ϵ per mode and travel price elasticities σ per mode are used to directly determine TD. The elasticities state that a percentual change in income I ($/cap/yr) or travel price TP ($/cap/pkm) causes a percentual change in TD. For example, with an ϵ of 0.5 a 10% increase in income results in a 3.1% increase in TD. The model consists of a single equation:

$$TD_{r,t,m} = TD_{r,t-1,m}\left(\frac{TP_{r,t,m}}{TP_{r,t-1,m}}\right)^{\epsilon_m}\left(\frac{I_{r,t}}{I_{r,t-1}}\right)^{\sigma_m}\left[\frac{pkm}{cap\ yr}\right] \quad (2)$$

Suffices here are for region r, time t (years), and mode m, the dimensions mentioned earlier. This simple model can be used to make forward projections on the basis of published estimates of income- and travel price elasticities. Note that in this formulation price can at best be a very aggregate measure of the actual travel cost of any particular transport mode – for instance, gasoline price for cars and official tariffs for trains. A graphical overview of the model is given in Figure 7.

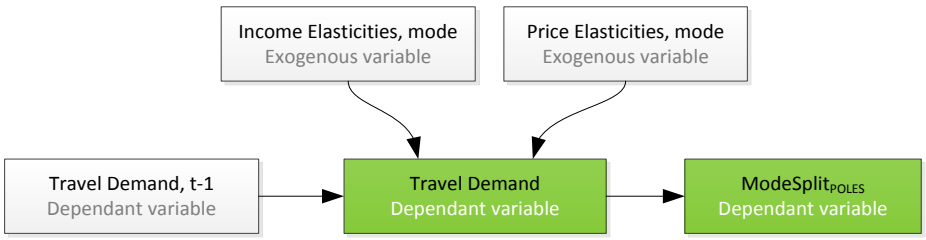


Figure 7: Simplified overview of the SIMPLE model

The values for σ and ϵ are given in Table 6. The values are derived largely from investigations in the UK, US and Europe. For PT the values vary between positive and negative depending on the source in literature, positive is decided on based on historic data. Details on the sources of the values can be found in the appendix to this thesis.

Independent variable	Symbol	Literature research value		Source
Travel price elasticity	σ	UM:	Unknown[4]	See appendix for full listing of sources (n~50)
		PT:	-0.8± 0.2	
		LDV:	-0.8± 0.2[5]	
		HSM:	-0.85 ± 0.2[5]	
Income elasticity	ϵ	UM:	Unknown[4]	See appendix for full listing of sources (n~50)
		PT:	-0.7 ± 0.3 / 1 ± 0.2[10]	
		LDV:	1.0 ± 0.3	
		HSM:	2.0 ± 0.5[5]	

Table 6: Values for the independent variables in the SIMPLE model

2.2.2 THE POLES TRAVEL SUBMODEL

The POLES model is similar to the SIMPLE model and adds to this concept by adding several modulators, including linear time trends TT, saturations SAT, and using equipment rates ER instead of travel demand for the growth of travel demand for LDV. Travel demand depends on elasticities on the Fuel Price FP ($/pkm), unlike the elasticity of the travel price used in SIMPLE. It additionally uses income elasticities ϵ on equipment rate ϵER. Moreover it uses fuel price elasticities δ instead of travel price elasticities. δ affects two years of history[6], using a weighing of 2/3rd and 1/3rd for the 1st and 2nd year, respectively.

The model has different equations for each mode, it's not symmetric in this sense. A generic equation can however be formulated. This formula can be filled in with the mode specific variables using Table 7. Notice that some of the factors are not present in every model, they become 1 or 0 and hence no longer affect the outcome.

$$TD_{r,t,m} = TD_{r,t-1,m}\left(\frac{FP_{r,t,m}}{FP_{r,t-1,m}}\right)^{\frac{2}{3}\delta_{r,m}}\left(\frac{FP_{r,t-1,m}}{FP_{r,t-2,m}}\right)^{\frac{1}{3}\delta_{r,m}} \left(\frac{I_{r,t}}{I_{r,t-1}}\right)^{\epsilon_{r,m}}\left(\frac{I_{r,t}}{I_{r,t-1}}\right)^{ER\epsilon_{r,t,car}}\left(1+TT_{r,m}\right)\left[\frac{pkm}{cap\ yr}\right] \quad (3)$$

[4] No literature for non-motorized has been found, hence the value is inferred from model calibration.
[5] For these modes no ranges have been found, hence their variance across sources has been taken.
[6] Notice that for the HSM mode the influence of price elasticity is over only a single year, it has been left out for simplification. This is a minor difference with this dataset. The full formulation can be found in the appendix.

The saturations mentioned work directly on the elasticities by reducing them from a starting income elasticity $\epsilon 0$ to zero near the saturation limit SAT. Thus stopping yearly growth as a function of income growth, because $x^0 = 1$. Notice that at saturation the TD growth is still effected by ϵ, and hence growing fuel prices. The factor g determines the strength of the saturation and has a value of 0.05[7].

$$\epsilon_{r,m,t} = \epsilon 0_{r,m} * \left(\frac{f_{r,\mathrm{t},m}}{f_{r,t-1,m} - SAT_{r,m}} \right)^{\mathrm{g}} [-] \qquad (4)$$

Table 7 convert the general formulations above to mode specific variants. All modes have a dependency on ϵ, δ and TT. The income elasticity ϵER of LDV works through the equipment rate ER rather than the travel demand. The saturation is based on a maximum on either equipment rate ERSAT (vehicles/cap) or on travel demand TDSAT (pkm/cap/yr) for a given mode, for LDV and HSM, respectively. PT does not have any saturation effects.

Notice that the fuel price FP for PT and LDV is the same. Initially the gasoline price for cars, later changing fuel types to biomass, hybrid and electric. The elasticity that works on FP is negative for LDV as earlier but positive for PT. This means that if fuel prices increase people tend to switch from LDV to PT.

Variable	Symbol	PT	LDV	HSM
Fuel Price	**FP**	$\mathrm{FP_{LDV}}$	$\mathrm{FP_{LDV}}$	$\mathrm{FP_{HSM}}$
Fuel Price Elasticity	$\boldsymbol{\delta}$	δ_{PT}	δ_{LDV}	δ_{HSM}
Income Elasticity	$\boldsymbol{\epsilon}$	ϵ_{PT}	0	ϵ_{HSM}
Income on ER Elasticity	$\boldsymbol{ER\epsilon}$	0	$\mathrm{ER}\epsilon_{\mathrm{LDV}}$	0
Time Trend	$\boldsymbol{TT}$	$\mathrm{TT_{PT}}$	$\mathrm{TT_{LDV}}$	$\mathrm{TT_{HSM}}$
Placeholder	$\boldsymbol{f}$	1	$\mathrm{ER_{LDV}}$	$\mathrm{TD_{HSM}}$
Saturation	$\boldsymbol{SAT}$	0	$\mathrm{ERSAT_{LDV}}$	$\mathrm{TDSAT_{HSM}}$

Table 7: Mode-specific values and variables for equations (3)(4) per mode. POLES

A graphical overview of the model is given in Figure 8.

[7] The authors only give an explicitly value for the LDV mode, the HSM mode uses the same equation but the actual value of g is not described. Latter is assumed to be 0.05.

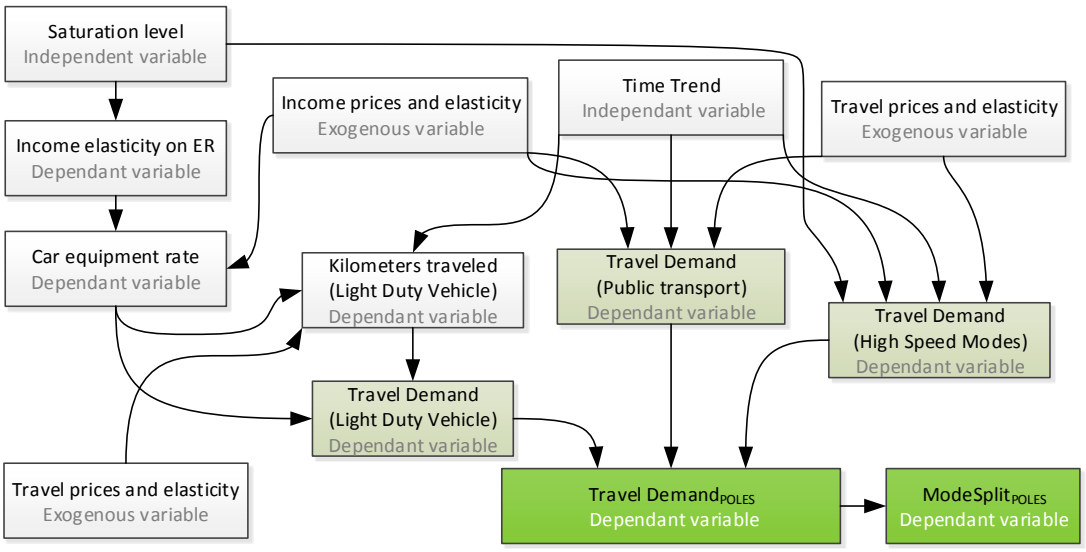


Figure 8: Simplified overview of the POLES model

Values for the POLES model were acquired through literature research and can be seen in Table 8 a full description can be found in the appendix. Values from the author were largely not available. Also notice that literature research on the transportation elasticities frequently limits itself to regional studies in OECD countries. It is therefore certainly not possible to determine the variables for each region, the databases that exists with this information are incomplete (Dahl, 2011) especially regarding developing countries. However, some information exists on the difference between developing and industrialized nations (IETA, 2007) and the development of elasticities using econometrical methods independent of country (Chamon, M. Mauro, P. Okawa, 2008; Dargay, Gately, & Sommer, 2007) but not enough information exists to do this for every variable, or for every region. This would entail 11x10=110 separate independent variables. The values are therefore based on investigations in the UK, US and Europe[8].

[8] Notice that these restrictions on data availability will also have affected the original authors and the set is therefore not unreasonable.

Independent variable	Symbol	Literature research value	Source
Fuel price elasticity	δ	PT[9]: 0.2±0.1	See appendix for full listing (n=50)
		LDV: -0.4± 0.2	
		HSM: -0.4 ± 0.2	
Income elasticity	ϵ	PT: -0.6 ± 0.3[10]	See appendix for full listing (n=50)
		LDV: 1.0 ± 0.3	
		HSM: 2.0 ± 0.5[5]	
Travel demand saturation level	TDSAT	HSM: Based on Schafer projections	(Schafer, Heywood, Jacoby, 2010)
Time trend	TT	Calibrated per mode, restricted range	(A. Kitous, 2000)
Income elasticity on equipment rate	ERϵ	LDV: 1.0 ± 0.25	(Chamon, M. Mauro, P. Okawa, 2008; Dargay et al., 2007)
ER Saturation level	ERSAT	Varies per region, based on econometric research (US ≈ 0.85 cars/capita).	(Chamon, M. Mauro, P. Okawa, 2008; Dargay et al., 2007)

Table 8: Values for the independent variables in the POLES model

2.2.3 THE GCAM TRAVEL SUBMODEL

The derivation of travel demand is similar to the formulation SIMPLE, except GCAM does it for all modes combined whereas SIMPLE does it per mode. In addition there is a calibration variable[11] s_r. Also notice that TD is derived for all modes combined, the elasticities[12] hence affect the aggregated TD. The fuel price elasticity σ works on the mode cost MC instead of the fuel price or travel price in the other models. The TD is later split up using the mode split MS using a price-based mode competition.

$$TD_{r,t} = s_r * \left(\frac{I_{r,t}}{I_{r,1971}}\right)^{\epsilon} * \left(\frac{\sum_{m=1}^{M} MS_{r,t,m} MC_{r,t,m}}{\sum_{m=1}^{M} MS_{r,1971,m} MC_{r,1971,m}}\right)^{\sigma} \left[\frac{pkm}{cap\ yr}\right] \quad (5)$$

Here the MC is averaged across all modes M by summing over m=1…M, weighed by the MS. The MS is a determined through a logit formulation, (Louviere & Woodworth, 1983; Yáñez, Raveau,

[9] The PT elasticity works on the fuel price of the LDV mode, not the PT mode, it is therefore positive

[10] Literature indicates a negative income elasticity, but this is not the case, as indicated by the authors. Additionally even within Europe this value changes significantly per region, e.g. from -0.05(FR) to -0.95 (UK) (Litman, 2012b). This is therefore the general elasticity for consumer products, approximately 1, as defined by Schafer (Schafer & Victor, 2000)

[11] Notice that this variable in fact gives the formulation a greater ability to fit to historic data when optimized

[12] Elasticities in the GCAM model are not the traditional elasticities in the sense that they work on the first (1971) year of the model simulation instead of the year t-1. This means in a strict sense they are not elasticities. For the purposes of this investigation, where projected datasets on income and price are generally increasing at a steady rate this does not make a large difference.

& Ortúzar, 2010). The share weight SW is a calibration factor which determines consumer preferences based on historic data. MC hence directly determines MS for each mode. The strength of the coupling between the MC and the MS is determined by the cost distribution factor μ, which is a constant.

$$MS_{r,t,m} = \frac{SW_{r,m} * MC_{r,t,m}^{\mu}}{\sum_{m=1}^{M} SW_{r,m} MC_{r,t,m}^{\mu}} [-] \quad (6)$$

The MC is finally determined by a combination of travel price and the value of time. The value of time is determined by dividing the wage rate WR ($/hr) by the door to door speed DDS (km/hr). The former is determined by dividing the income I ($/yr) by the hours worked HW (hr/yr). This assumes that every hour spent traveling is an hour of productivity and hence income lost.

$$MC_{r,t,m} = TP_{r,t,m} + \frac{\mathrm{WR}_{r,t}}{DDS_{r,t,m}} \left[\frac{\$}{pkm}\right] \quad (7)$$

A simplified overview of the model is given in Figure 9.

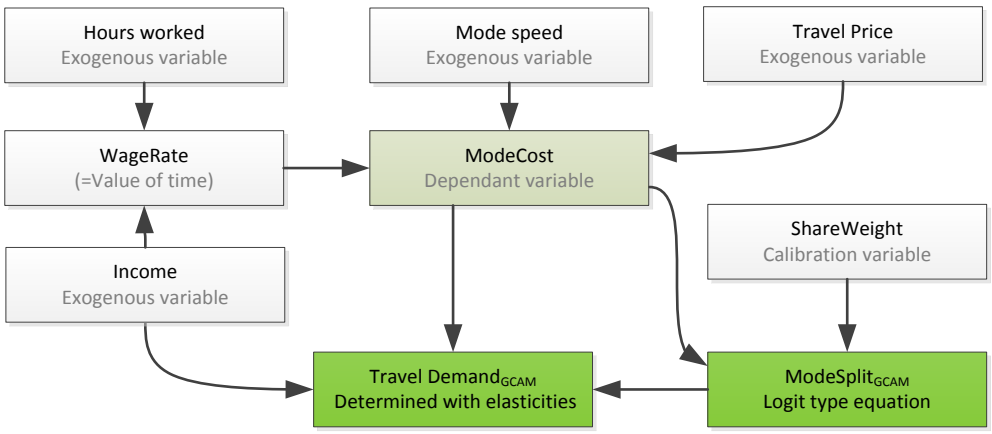


Figure 9: Simplified overview of the GCAM model

Parameter values for GCAM are known due to contact with the authors, availability of the original model and availability of input data. The values used in the GCAM model fit with other literature research, although they lack any variation across regions or modes. Therefore they are only valid at an aggregated level. The values are displayed in Table 9.

Independent variable	Symbol	Value according to author	Literature research	Source
Cost distribution Parameter	μ	-2	N.A.	Digitally provided by (Kyle & Kim, 2011)
ShareWeight	SW	Determined by calibration (LDV = 1)	N.A.	(Kyle & Kim, 2011)
Price Elasticity	σ	-1 to -0.8	-0.8 ± 0.2	See appendix for full listing (n~50)
Income Elasticity	ϵ	1	1 ± 0.3	See appendix for full listing (n~50)
Sigma	s	Calculated from TD, I and TP	N.A.	(Kyle & Kim, 2011)

Table 9: Values for the independent variables in the GCAM model

2.2.4 THE TIMER TRAVEL SUBMODEL

The TIMER travel submodel is quite different from the other models presented in that it does not have a travel demand dependent on pricing and income through elasticities. Instead the total TD across all modes is determined using the constraints on travel money budget TMB (% of income) and travel time budget TTB (hr/yr). The TD hence calculated, is split up using price competition similar to the formulation of GCAM[13].

$$TD_{r,t,m} = \frac{MS_{r,t,m}\, I_{r,t}\, TMB_{r,t}}{\sum_{m=1}^{modes}(MS_{r,t,m}\, TP_{r,t,m})} \left[\frac{pkm}{cap\ yr}\right] \quad (8)$$

MS is calculated using a multi-nominal logit (MNL) formulation. The concept is very similar to that of the GCAM competition, but no identical. The normalized prices MCR determine the mode share and are modulated through the factor λ, which determines the strength of the interaction. The change in MS over time $\partial MS/\partial t$ can be limited through the inertia α which is used to factor in infrastructural change.

$$MS_{r,t,m} = \alpha \frac{e^{-\lambda\, MCR_{r,t,m}}}{\sum_{m=1}^{modes} e^{-\lambda\, MCR_{r,t,m}}} + \left(MS_{r,t-1,m} - \alpha\, MS_{r,t-1,m}\right)[-] \quad (9)$$

The normalized mode cost MCR is calculated by combining two factors. On the left hand side TP[14] and the preference factor PF, which is a calibration factor modelling consumer preference in a similar way GCAM does. PF is determined from two independent calibration variables which are calibrated separately, $PF_{r,t,m} = PFStart_{r,m} + t\, PFSlope_{r,m}$ PFstart determines the starting

[13] The competition in GCAM crucially differs in time evolution

[14] TP is modified by the comfort elasticity, which is taken as a factor in the vehicles module, this is described more generally in the datasets section for it applies generally to all models. It is taken into the monte carlo simulation only for TIMER-travel.

value of PF and PFslope the yearly increase. On the right hand side the door to door speed DDS (km/hr) and time weight TW, which constrains the overall time use TU (hr/yr) to the TTB. Both sides are normalized using the averaged TP and DDS, respectively.

$$MCR_{r,t,m} = \frac{TP_{r,t,m}}{PF_{r,t} \sum_{m=1}^{M} MS_{r,t,m}\, TP_{r,t,m}} + \frac{TW_{r,t}\ \sum_{m=1}^{M} MS_{r,t,m}\, DDS_{r,t,m}}{DDS_{r,t,m}}\ [-] \qquad (10)$$

TTB is taken to be constant over time and regions and is a constrained on the model it ensures the TU remains near the TTB (TU ≈ TTB). This works through use of the time weight TW. The TW works on the normalized mode cost MCR which in turn affects the MS and the travel demand, changing the time use. The TW is adjusted every year in order to fulfill this requirement. There are two methods of doing this, iteratively and through optimization. They were published in separate publications of the model, (Bastien Girod et al., 2012) and (B. Girod, Van Vuuren, & De Vries, 2013), respectively.

$$\text{1) Iterative approach: } TW_{r,t,m} = \frac{TW_{r,t-1} * TD_{r,t,m}}{TTB\ * DDS_{r,t,m}}\ [-] \qquad (11)$$

$$\text{2) Minimizing approach: } min\left(\left(TTB - \frac{TD(TW)_{r,t,m,}}{DDS_{r,t,m}}\right)^2\right) \text{ by changing } TW \qquad (12)$$

The TMB is a constraint on the total travel cost TTC (% of I), it keeps the TTC exactly equal to the TMB. The TMB without additional factors increases from 3.5% to 12% with development of the MS. It is calculated by summing over the MS of LDV and HSM. It is corrected regionally by a correction factor CF which is determined econometrically. TMBA is a variable used for sensitivity analyses.

$$TMB_{r,t} = 0.035 + 0.085 \sum_{\text{m≥car}}^{\text{modes}} \text{MS}_{\text{r,m,t}} + CF_r + TMBA[-] \qquad (13)$$

A simplified overview is given in Figure 10.

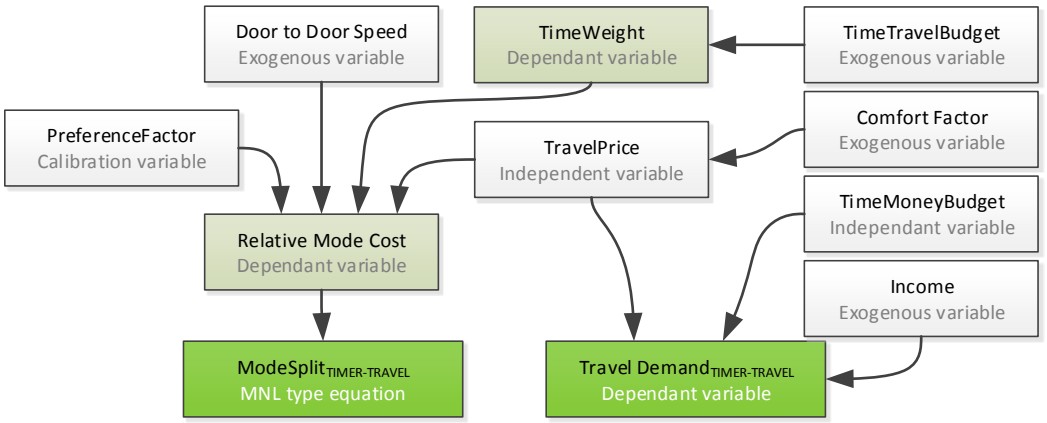


Figure 10: Simplified overview of the TIMER-travel model

The values used by in TIMER-travel are known, for they been discussed with the author. Nonetheless, literature research has been performed on all variables. Some have no direct foundation in literature and are exclusively determined by calibration. The values, descriptions and sources are given in Table 10.

Independent variable	Symbol	Value by author	Literature research	Source
Travel Time Budget	TTB	438 hr/yr	365 – 550 hr/yr	(Schafer, Heywood, Jacoby, 2010), (Avineri, 2009; Chen & Mokhtarian, 2008; Stopher & Zhang, 2011)
Travel Money Budget	TMB	3.5% to 12%, increasing with car ownership, corrected per country	3.5% ± 2% to 12% ± 3%,	(Schafer, Heywood, Jacoby, 2010; Stopher & Zhang, 2011)
Substitution Elasticity	λ	3.2	N.A.	(Bastien Girod et al., 2012)
Preference factor Start	PFstart	0.01-100, determined by calibration	N.A.	(Bastien Girod et al., 2012)
Preference Factor Slope	PFslope	0-0.1, determined by calibration	N.A.	(Bastien Girod et al., 2012)
Comfort Elasticity	ψ	0.5	0.5	(Bastien Girod et al., 2012)
Inertia	α	1, assumed redundant	N.A.	(Bastien Girod et al., 2012)

Table 10: Values for the independent variables in the TIMER-travel model

2.3 CONVERGENCE TO INDUSTRIALIZED REGION

Convergence of consumer preferences and hence the travel demand per capita for a certain travel mode is a feasible future scenario. POLES is the only model presented that natively implements this through saturations. As discussed in section 4 one of the main issues with the present projections is developing regions are at an early stage in developing their mobility. This results in preferences that are different from the industrialized regions for several reasons, among them the small datasets, early development and possible different wealth distribution among the population. This issue is despite projections in range of literature results, replicating author results, replication and a relatively objective methodology.

It is therefore decided to modify the GCAM and TIMER-travel models in order to allow the readers to judge the issues themselves. The models are hence enhanced beyond literature definitions, and have an effect added that converges them to industrialized consumer preferences in the future. These results and the original models will be presented, because it is impossible to objectively state which is better. For this would be based on the forecasts themselves not their performance in the past, which is unimpeded by this modification.

For both TIMER-travel and GCAM, convergence is not natively implemented. POLES assumes pre-set convergence levels and adjusts elasticities accordingly. SIMPLE doesn't assume any regional differences or convergences. Because both GCAM and TIMER-travel have a similar notation with regards to consumer preferences, share weight and Preference factor respectively. The reader can see in the individual analysis section (2)) that both GCAM and TIMER-travel over represent HSM in the early calibration, when compared to the first world.

The variables are found to be strongly dependent on income, as depicted in Figure 11 for TIMER-travel and Figure 12 for GCAM. The values displayed are for the best fitting runs used for the Monte Carlo simulations used in section 0. The right graph is defined as the inverse to the power of two in order to make it compatible to the Preference factor and show a linear relationship. Additionally the ratio of LDV and HSM is used, this has to do with a difference in the way variables are calibrated in the two models, because of the difference in formulation. For TIMER-travel this tends to maximize NM, which serves as a reference, in GCAM the values are all relative.

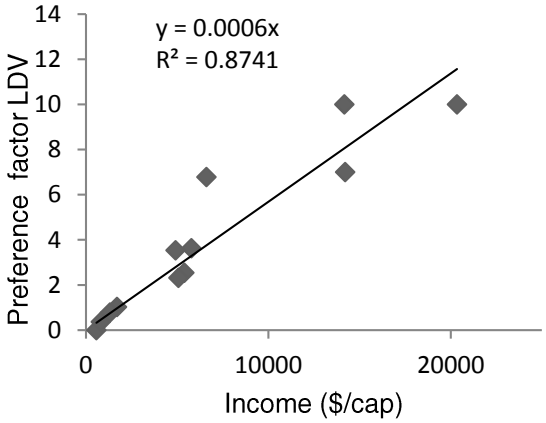


Figure 11: PF LDV per region as a function of I in the year 2005 for the TIMER-travel model. A simple regression is used to fit the line

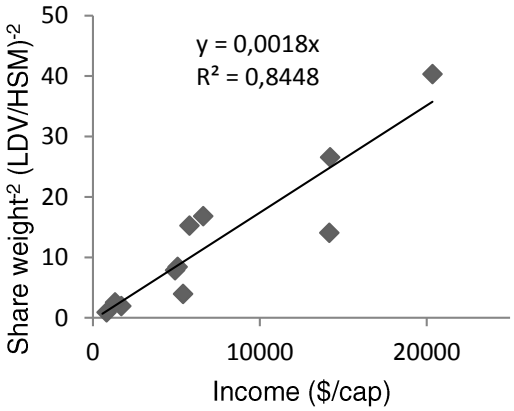


Figure 12: SW of LDV relative to SW of HSM per region in the year 2005 for the GCAM model, the ratio is taken to the power minus two. A simple regression is used to fit the line

There is hence an indirect relationship between both share weight SW and Preference factor PF with respect to regional income I ($/cap/yr), as depicted in the graphs above. A method can be applied that allows these two variables to converge to the values of the industrialized region, as these regions are the most reliable references of (economic) development to date. The following modifications are used for this convergence for Timer-Travel:

(TIMER-travel)

$$PF_{r,t,LDV} = \frac{I_{r,t}}{\overline{I_{Ind,t}}} * \overline{PF_{Ind,t0,LDV}} \; [-] \qquad (14)$$

The subscript Ind and bar denotes the population weighted average of the industrialized nations in both the I and PF variables. Population averaging is done using $\bar{x} = \sum_{r=1..R} x_r \, Pop_r$ with R the number of regions and Pop the population per region. This formula thus recalculates the PF each year, based on the average PF of industrialized nations (calculated with a regression, as seen in the graphs above) and the ratio of the regions income to that of the average industrialized nation. The normal calibration for PFstart and PFslope apply.

The GCAM model has a similar formulation, but has to be corrected for the fact that the definition of SW is inverse to that of PF and the squared taken. Therefore some rewriting is necessary.

(GCAM)

$$SW_{r,t,LDV} = \overline{SW_{r,t,HSM}} \frac{\sqrt{\overline{I_{Ind,t}}}}{\sqrt{I_{r,t}}} \frac{1}{SW_{Ind}} \; [-] \qquad (15)$$

The subscript ind again denotes the average of industrialized nation. This variable is determined from the regression in Figure 12 to be between 35 and 40. Applying this method yields the results presented in Figure 13 for TIMER-travel and Figure 14 for GCAM.

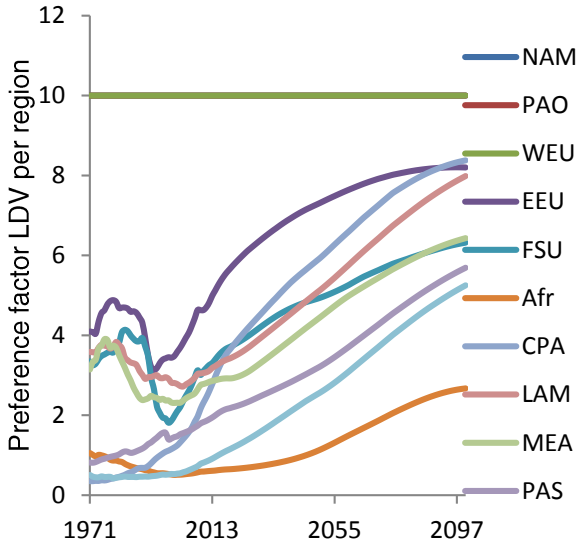


Figure 13: Preference factor, TIMER-travel model, over time for each region, convergence can be seen toward the industrialized region

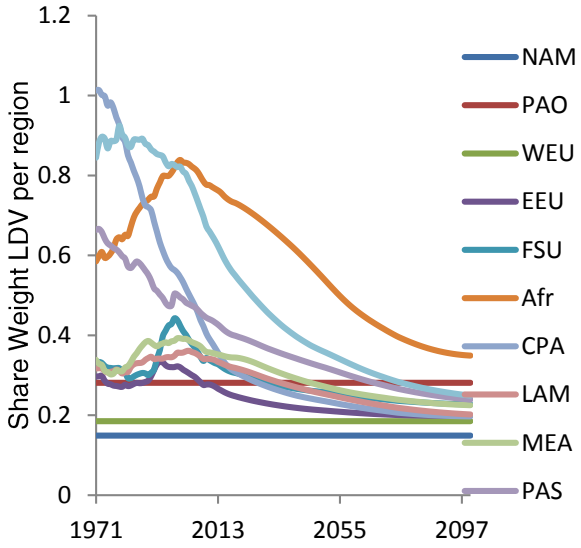


Figure 14: Share weight, GCAM model, over time for each region, convergence can be seen toward the industrialized region

2.4 DATASETS / EXOGENOUS VARIABLES

This section concerns the datasets used for all values that do not have probabilistic ranges, the exogenous variables. These values are of the independent variables described in the previous section. The section is split into four sections, first describing the datasets taken directly from various source, then the datasets that have been modified, finally datasets constructed from mathematical formulations, mostly based on econometric analyses.

2.4.1 ORIGINAL DATASETS

In order to make a sensible projection of the future the models require both datasets for calibration and datasets that set scenarios for the future. A summary of the exogenous variables and their sources are described in Table 5. For the remained the full plots can be found in the appendix or online on http://globaltravelmodels.codeplex.com. The emission scenarios are given more explanation.

The datasets were originally derived for 26 regions and 7 modes of transportation. In order to accommodate all the models they have been recalculated to 11 regions and 4 modes. This recalculation has been done based on the units of each dataset using simple matrix operations. For example to recalculate the average income ($/cap) for a different set of regions the income is multiplied by the population of a given region then divided by the population of the new region. Sets that depend on the total pkm per year, the very indicator of this thesis, such as speed (pkm/hr) had to be recalculated from the bottom up using published methods (Bastien Girod, Deetman, & Van Vuuren, 2011). If the reader is interested in the details of the conversion they can be seen in the excel and Matlab files online.

Variable	Year range	Scenario	Source
Income	1971-2100	TIMER baseline scenario	(Bouwman et al., 2006)
Population	1971-2100	TIMER baseline scenario	(Bouwman et al., 2006)
Urban Population	1971-2100	TIMER baseline scenario	(Bouwman et al., 2006)
Speed[15]	1971-2100	TIMER-travel baseline	(Bastien Girod et al., 2012)
Speed (GCAM)[16]	1971-2100	Constant for all years, TIMER baseline at year 2005	(Bastien Girod et al., 2012)
Vehicle price	1950-2100	TIMER-travel original scenario	(Bastien Girod et al., 2011)
Hours worked[17]	2005	Constant over time	(Laborsta, 2012)

[15] Vehicle speed for non-motorized, public transport and cars is constant and depends on the percentage of population in urban areas. For Trains and airplanes an increasing trend is used up to 2100

[16] The GCAM paper keeps the speed of all modes constant, TIMER-travel increases the speed of air significantly, the model does not take this into account in a way that prevents air from exploding. After contacting the authors of the paper it's decided the speed has to remain constant.

[17] Hours worked data is sparse for non-OECD regions, for them the most recent valid year has been used. No reliable data exist for Africa and 2000 hours/years is used as an estimate

Load factor	1971-2100	Projections based on income, see Figure 15 and Figure 16	(Airbus, 2012), (Federal Aviation Authority, 2012), and (Fulton & Eads, 2004)
Energy price	1971-2100	Reconstructed TIMER baseline, see vehicles module	(Bouwman et al., 2006)
Non-energy price	1971-2100	Reconstructed TIMER baseline, See vehicles module	(Bastien Girod et al., 2012)
Gasoline price	1971-2100	TIMER baseline	(Bouwman et al., 2006)
TD	1950-2005	Schafer data	(Schafer, Heywood, Jacoby, 2010)
Travel price	1971-2100	Reconstructed TIMER baseline, See vehicles module	(Bastien Girod et al., 2012)
Equipment rate	1971-2100	Reconstructed econometric model	(Chamon, M. Mauro, P. Okawa, 2008; Dargay et al., 2007)
ER saturation	∞	Reconstructed econometric model	(Chamon, M. Mauro, P. Okawa, 2008; Dargay et al., 2007)
TMB correction	1971-2100	Based on travel data from Schafer	(Schafer, Heywood, Jacoby, 2010)
Energy intensity	1971-2100	Based on TIMER data	(Bouwman et al., 2006)
CO_2 intensity	1971-2100	Based on TIMER data	(Bouwman et al., 2006)
RCP scenarios	2000-2100	Used for CO_2 scenarios	(Bastien Girod, Van Vuuren, & Hertwich, 2013)

Table 11: Datasets and sources for the exogenous variables, all datasets can be found in excel format on http://globaltravelmodels.codeplex.com

2.4.2 MODIFIED DATASETS

In order to simulate changes in the Comfort Factor used in TIMER-travel the travel price is produced by the, simplified, vehicles module of the TIMER-travel model. The vehicles module is not reconstructed entirely, but it is derived from previous model runs. For the POLES and SIMPLE models, however, the travel price is not explicitly defined and therefore the TIMER-travel baseline is assumed with a comfort elasticity ψ of 0.5 (Bastien Girod et al., 2012). The travel price is split into three cost components, the non-energy costs NEC ($/vkm), the energy costs EC ($/vkm) and the other costs OC ($/vkm). The load factor LF (pkm/vkm) is used in order to convert from the vehicle to the person. The energy cost is derived using an equation dependent on the net present value, energy efficiency and fuel costs (Bastien Girod, Van Vuuren, et al., 2013) not presented here.

$$TP_{r,t,m} = \frac{NEC_{r,t,m} + OC_{r,t,m} + EC_{r,t,m}}{LF_{r,t,m}} \left[\frac{\$}{pkm}\right] \quad (16)$$

In order to include the ψ as a variable in the Monte Carlo simulations the existing dataset is modified for each value of ψ, rather than recalculating the entire dataset from the ground NEC is recalculated by using published references values for the NEC (in the USA, 2000) (Bastien Girod

et al., 2012). Recalculating would lead to different values than those published. Because the ψ only impacts the NEC at time t it is decided to formulate NEC as a function of the ratio $\psi_{new} = 0.5$ to ψ_{old}.

$$NEC_{r,t,m,new} = NEC_{r,t,mold} \frac{\alpha_{r,new}}{\alpha_{r,old}} \left(I_{r,t,m}\right)^{\psi_{new} - \psi_{old}} \quad (17)$$

Notice that here $\alpha = I_0^{-\psi}$ is a constant[18]. The new and old subscripts denote the present model run and the data available from the past model, respectively. The energy cost, other cost and non-energy costs are recombined in order to produce the travel price in the scenario with differentψ.

The GCAM model has a different approach to the mode cost than the TIMER-travel formulation. The income dependency of the formulation is included not through a ψ, section 2.2.1, as in Timer Travel but comes though the wage rate, see section 2.2.3, which affects the price. The TIMER-travel dataset with this income effect would hence make the price doubly sensitive to income. It is therefore removed simply by setting ψ_{new} to ψ_{old}. This sets the NEC to the same level at every point in time.

The reconstructed load factors for LDV and PT are given in Figure 15 and Figure 16. These load factors are modeled as a power function fitted to income and are hence region and time dependent. The load factors for air travel are based on a literature investigation and extrapolation of the most reliable forecasts (Airbus, 2012; Federal Aviation Authority, 2012), for the used capacity (% of used seats) and total capacity (seats per airplane). Extrapolating these using a simple power regression gives the expected load factor. For the last mode, non-motorized, the load factor is unity, for only one person is assumed to walk or use a bike at any one time.

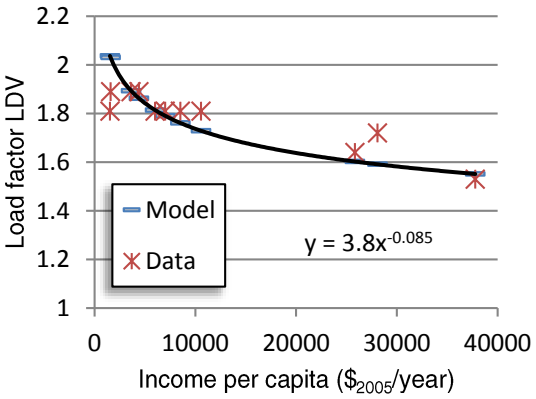


Figure 15: Modeled load factor compared to actual data for LDV (Fulton & Eads, 2004)

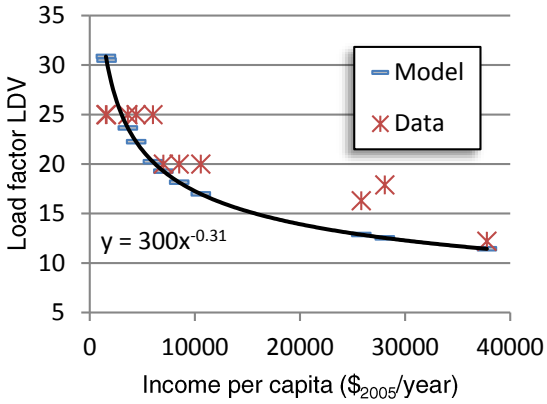


Figure 16: Modeled load factor compared to actual data for PT (Fulton & Eads, 2004)

The same literature investigation is also performed on the load factor for planes (J. J. Lee, 2000), as depicted in Figure 18. Calculating this load factor is complicated by the fact that aviation literature splits the variables up in seat load factor (Airbus, 2012), and the total number of seats

[18] Also notice that this alpha is not the same one used earlier to denote the inertia

(Federal Aviation Authority, 2012; J. J. Lee, 2000). Both are extrapolated beyond the horizons in the papers, which only forecast up to 2032. The load factor in terms of passengers per plane is next calculated based on these variables.

The resulting dependency of the forecast on the ψ is depicted in Figure 18: Travel price per mode, from top to bottom, HSM, LDV, PT, UM, for various comfort elasticities ψ (0.5 is the baseline scenario) Figure 18. It can be seen that the travel prices decrease with decreasing ψ.

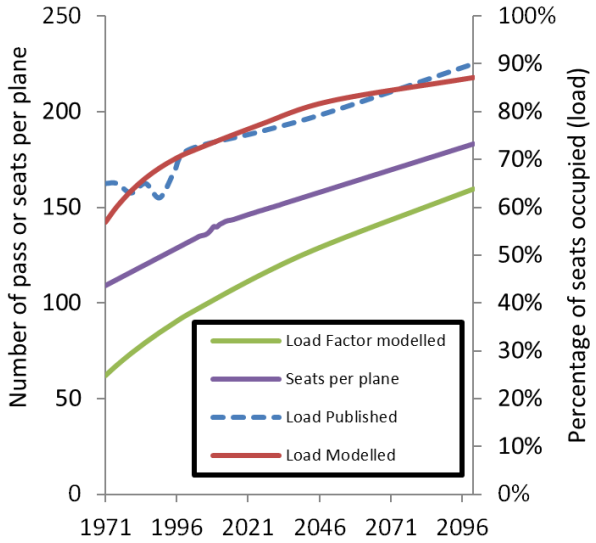


Figure 17: Construction of the modelled load factor based on the projected growth of seats and load. The modelled load factor is fitted to the publishied load factor hence constructed

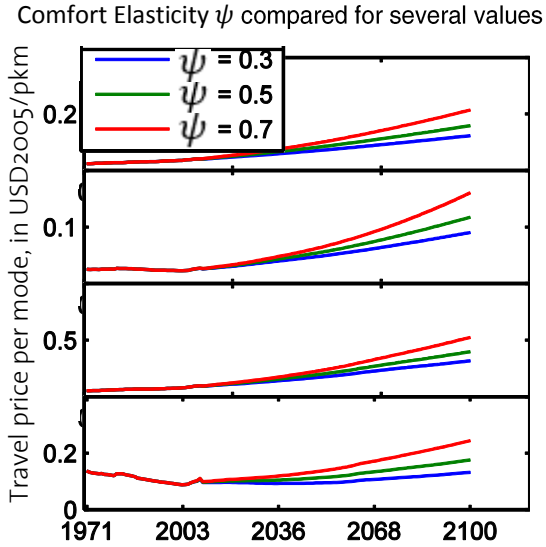


Figure 18: Travel price per mode, from top to bottom, HSM, LDV, PT, UM, for various comfort elasticities ψ (0.5 is the baseline scenario)

2.4.3 CONSTRUCTED DATASETS

For the POLES model effort has been put toward the acquisition of equipment rates data and saturation levels. The saturation levels for cars have been acquired from a publication that has a convincing econometric model for their development of the publications investigated (Chamon, M. Mauro, P. Okawa, 2008). Based on this model the saturation levels have been calculated per region, Figure 20. An overview of the forecasted development as a function of time is given in Figure 19. The present-day values are determined using the initial values of the same paper.

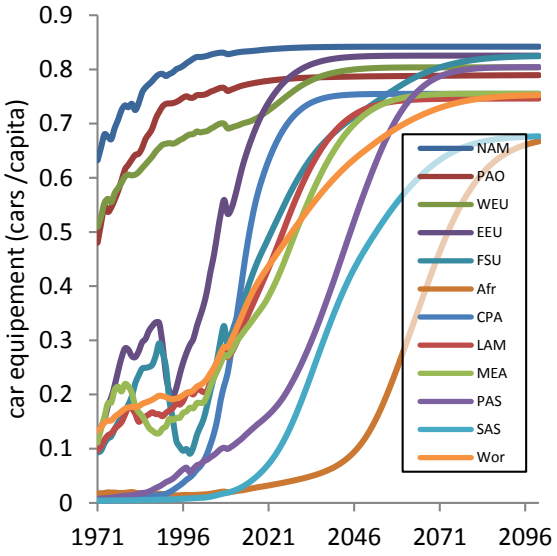


Figure 19: Historic data (<2006) and projection (>2005) of per capita equipment rate of cars per region (Chamon, M. Mauro, P. Okawa, 2008)

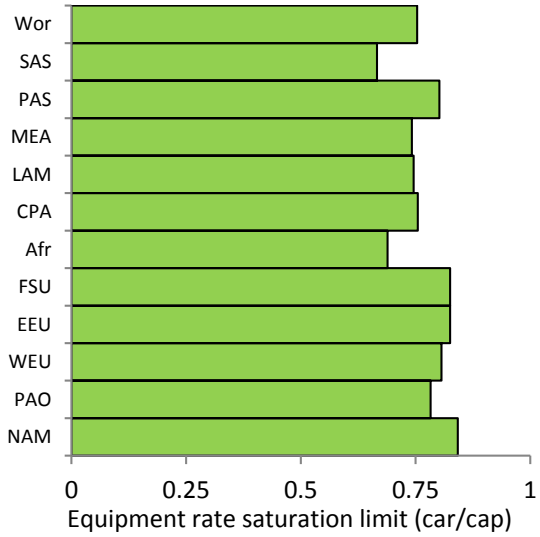


Figure 20: Saturation limit to per capita equipment rate of cars based on the asymptotic value of Figure 20

A similar investigation as used adobe is performed for the saturation levels of airplanes. This is largely based on a publication by Schafer (Schafer, Heywood, Jacoby, 2010). Although he explicitly mentiones the saturation level given infinite income (the target point, at 240.000 pkm/yr, if all time were spent in airplanes) his maximum estimates for the air demand in 2100 are instead used, and regionally dependent as depicted in Figure 21. Due to the fact that this cap is strict to the rapidly developing region, see the discussion on saturation in section 4, compared to industrialized, it is decided to equalize the values for developing region to that of the Industrialized region at 50000 pkm/yr, this is in line with an assumption of converging consumer preferences and hence modal choice (Schafer, Heywood, Jacoby, 2010).

For the GCAM model the wage rate has to be calculated, which is a function of the hours worked and the income (Keynes, 2007). The hours worked are assumed fixed as per the model (Kyle & Kim, 2011) in 2005 or whatever most recent year is available. The wage rate is assumed to increase linearly with income (Kyle & Kim, 2011). The data is derived from several databases, notably (Laborsta, 2012; United Nations, 2009). The calculated wage rate in 2005 is given in table Figure 22.

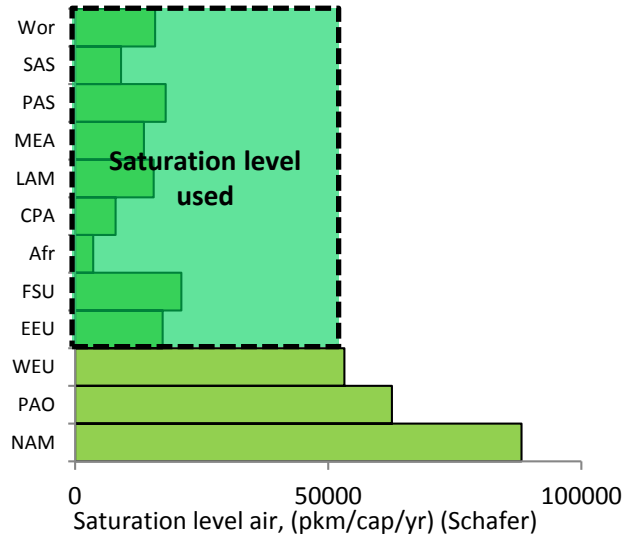


Figure 21: Saturation levels of air travel derived from Schafer, the dashed line represents the values used for developing regions

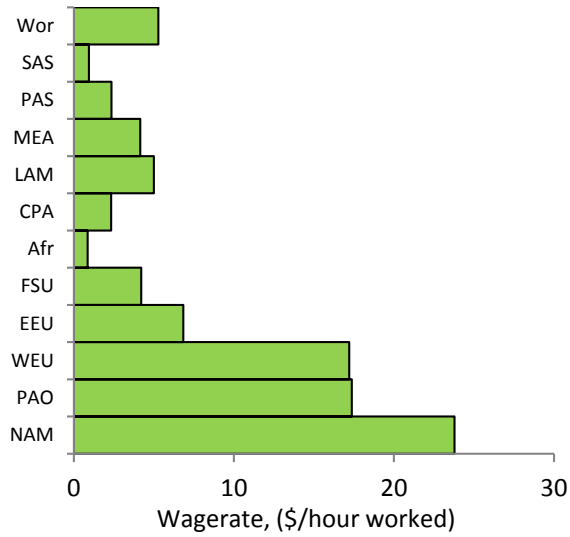


Figure 22: Wage rate per region in the year 2005

The value taking for the TMB variable used for the TIMER-travel model is corrected based on a regional bias. This is necessary to correct for various regional differences not directly caused by income, and has been performed by the original author (Bastien Girod et al., 2012). As defined in section 2.2.1. The correction factor (CF) is determined by a linear regressing over the years of historic data per region, comparing the actual TMB to the TMB calculated with the formula above and minimizing the squared residual by adjusting the CF. The resulting variables can be found in Figure 23.

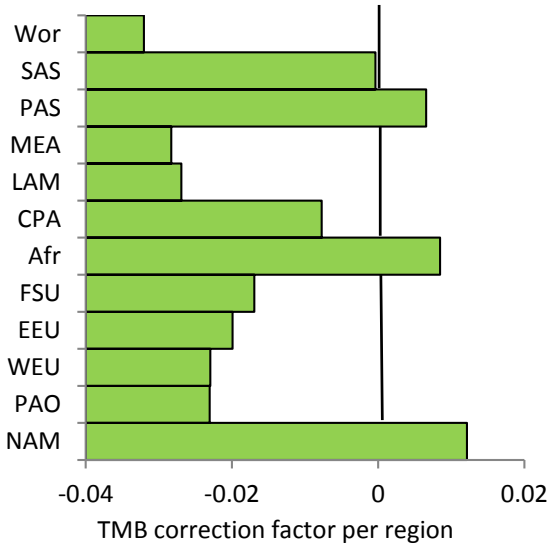


Figure 23: Correction factor CF of the TMB calculation for each region

2.5 CALIBRATION

Calibration takes place using manual calibration and automated calibration. The process used by manual calibration is used to determine the ranges of the variables in probabilistic projections and is described in section 0. The calibration discussed here involves the automatic calibration, also known as optimization or minimization. This section will first describe the optimization procedure and next the sensitivity analyses used to investigate the validity of the optimization.

2.5.1 OPTIMIZATION

Generally optimization problems rely on a single scalar objective function that is to be minimized or maximized. However, the type of optimization in the present investigation requires all four independent modes of the TD dataset be optimized. In general optimization allows for the reduction of only a single objective function, for example the squared difference between the actual dataset and the predicted dataset. The present problem is, however, a more complicated multiple objective optimization problem (Marler & Arora, 2004). Literally hundreds of solutions exist to these types of problems, each specific to certain types of problems.

The range of possible solutions created by having multiple dimensions to be minimized is set of pareto-optimal solutions (Modelling & Bilthoven, 1995). Some of these solutions might be more appropriate to the problem at hand than others, but there is no way of generally defining this for each problem. Time was invested to investigate the literature on the topic, this process and its references will not be discussed here. The reader is encouraged to read up on the topic in the excellent publication "Survey of multi-objective optimization methods for engineers" (Marler & Arora, 2004)

Many possible solutions to the multi-objectivity problem were attempted and it is decided to define a single objective function with weighing for each mode.

$$OF = \sqrt{\sum_{t=t_0}^{t_{end}} \frac{\sum_{modes} Weight_m \left(TD_{schafter,m} - TD_{model,m}\right)^2}{\sum_{modes} \left(TD_{Schafer}\right)^2}} \qquad (18)$$

With OF the objective functions, t_o and t_{end} the first and final years of the optimization, respectively, $Weight_m = \frac{\sqrt{TD_m}}{\sum_{modes}\sqrt{TD_m}}$, a weighing variable that means large modes are counted more heavily, but not linearly,$TD_{schafer}$ the historic TD according to (Schafer, Heywood, Jacoby, 2010) and TD_{model} the predicted TD by the model, see Figure 24. If more years are used for a given calibration run then the objective function is simply the average of the objective function depicted above. (Ruijven et al., 2009)

The solvers considered for the optimization are depicted in Table 12. Each of these solvers comes with upsides and downsides, and it is decided the genetic algorithm is used in combination with FminSearchbnd to find an optimal solution. The genetic algorithm does not require starting conditions and is stochastic, without any assumptions on starting values it allows for a global

solution to be found. The FminSearchbnd was constructed based on the observation that, if left unbounded, the TIMER-travel model would sometimes develop unrealistic values. This function is essentially FminSearch with a potential well added.

Name	Type	Notes
***Ga**	Global optimization	Used when no starting assumption are made
FminSearch	Unbound smooth local	Rejected after solutions out of boundary
FminCon	Bound local	Rejected when solution was insensitive
Lsnonlin	For Least Square problems	Required reformulation of problem
Fminbnd	Bound not smooth local	Rejected when solution has few decimals
***FminSearchbnd**	Bound smooth local	Adjusted version of Fminsearch, high sensitivity, resolution and manually added boundaries (i.e. steep walls)

Table 12: Optimization methods considered and tested *denotes the methods used

For the GCAM model the calibration is adjusted for better results, because the fit to the mode share is independent. And through the price mechanism the fit of the total TD does depend on the mode share. It is therefore decided to first optimize the mode share, then the total TD and finally the normal objective function. This method produced superior results and is used for this model.

2.5.2 SENSITIVITY ANALYSIS

The validity of the results depends on the sensitivity of outcomes to the independent calibration parameters and sensitivity of the output to the input datasets. The sensitivity analysis here is defined as a tool to investigate the effect of calibration and literature parameters on the model outcome, probabilistic projections are defined as using different scenarios and ranges on the input datasets and investigating their impact on output.

Sensitivity analysis will be used in order to determine the stability of the model under changes in the input variables provided by literature. A method called "one at a time" sensitivity analysis (Klepper, 1997) will be used to investigate sensitivity. First a model simulation is ran which determines the appropriate values for each independent variable. Each value is next normalized to 1, varied across a range 0.2 – 2 time the original value of the variable, and fed back into the model keeping all other variables equal. The impact on the TD in the year 2100 can then be plotted against the change in input (Crick, Hill, & Charles, 1987), see Figure 25. The non-normalized variable value refers to the original value now represented by one. For example μ was originally -2 and next normalized to 1 for easier comparison.

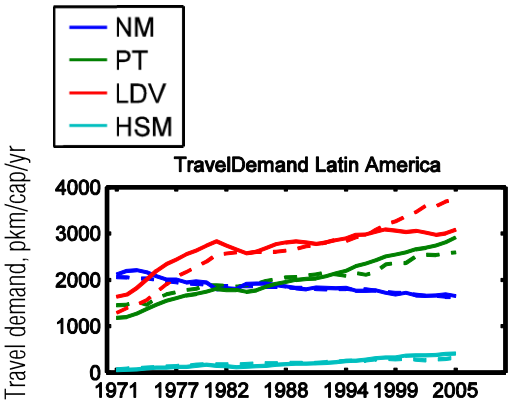


Figure 24: Example of a projection for four modes. Dashed lines indicate historic data, continuous lines indicate predicted data

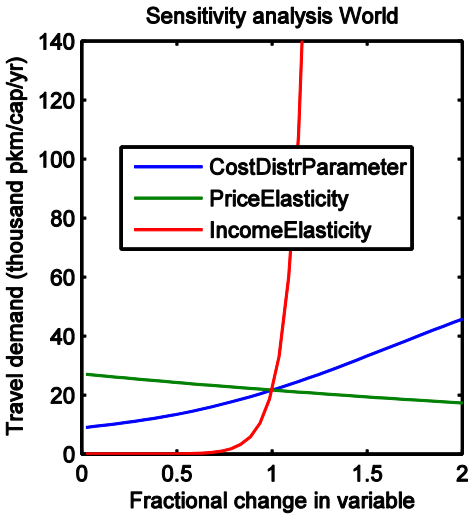


Figure 25: Example of a sensitivity analysis for GCAM. Non-normalized values are -2, 1 and 1 for μ, σ and ϵ, respectively

2.6 VALIDATION

In order to ensure a good historic fit of the model it is tested in predicting a known dataset, here using several historic years to forecast known historic data. The model is calibrated from 1971 to 1980, and allowed to project the 25 years up to 2005. First the error is investigated qualitatively, it is tested on behavioral soundness. Predicted and historic values for travel demand are hance plotted together, see Figure 24. It is left to the reader to accept or reject these samples for not all can be presented. To this end some of the best (Western Europe, World) and worst fitting (South Asia, Africa) are presented for each model. Second the error measures are taken quantitatively.

2.6.1 RESTRICTIONS IN TMB AND TTB

In order to further judge the validity of a model in a given region the travel time budget TTB and travel money budget TMB are used as empirical boundaries on the time use TU and total travel cost TTC, respectively. Both should be within bounds, $TTB_{min} < TU < TTB_{max}$ and $TMB_{min} < TTC < TMB_{max}$, given by Schafer and other authors(Schafer, Heywood, Jacoby, 2010). Notice that the TTB and TTB are empirical findings whose validity is disputed in literature, therefore exceeding them by a small amount could still be acceptable especially in a future scenario. However large deviations, such as a TU >24h/day are unacceptable. Notice that the TU sometimes takes into account all modes, and sometimes motorized modes only. Including only motorized lowers the lower bound to 0, because a underdeveloped society could have 100% NM and hence no motorized TU.

Two examples of the analysis used are given in Figure 39 and Figure 40. This analysis takes the best model run (i.e. with the lowest TDfit), and plots every yearly value for TU or TTC in a box plot. The figures therefore represent an evolution over time of the TTC and TU. The green box indicates the values allowed according to literature.

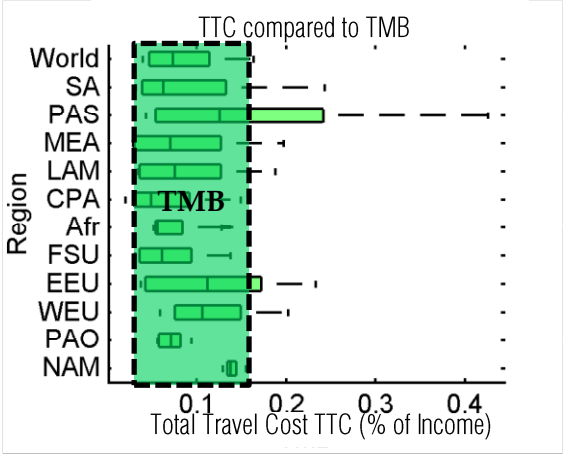


Figure 26: Boxplot of TTC for 130 years for the best Monte Carlo run of all regions. Dashed box represents the range of found values of TMB.

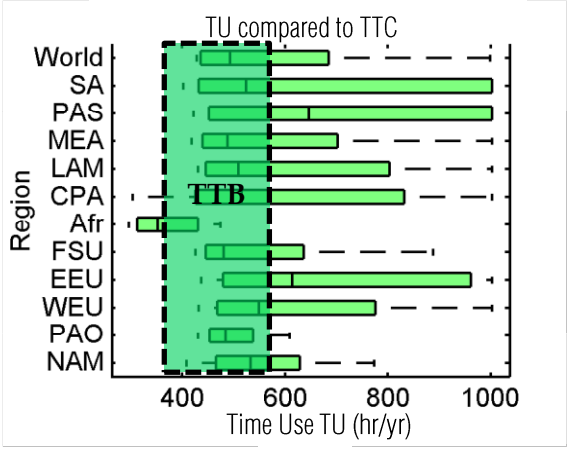


Figure 27: Boxplot of TU for 130 years, all modes, for the best Monte Carlo run of all regions. Dashed box represents the range of found values of TTB.

2.6.2 ERROR MEASURES

An error is indicated, based on the normalized root-mean squared error (NRSME) (Ruijven et al., 2009). The error indicates the average normalized difference between the total TD predicted and historic, it is called the TDFit. It is defined as follows:

$$TDFit = \frac{\sqrt{\sum_{t=1971}^{2005}\left(\frac{TD_{historic,t} - TD_{predicted,t}}{TD_{historic,t}}\right)^2}}{2005 - 1971} \qquad (19)$$

With $TD_{historic}$ the total historic TD and $TD_{predicted}$ the predicted TD. Notice that this error is not the same as the objective function, for in the objective function all modes are compared separately, whereas here they are compares as total TD. If a model proves valid in the behavioral sense, in that the historic projections follow the same behavior as data, the fit is taken as accurate if the error is below 10% (Ruijven et al., 2009).

2.7 DETERMINING RANGES, PROBABILISTIC PROJECTIONS

One of the methods of model comparison used in the present investigation is by using probabilistic projections. This section explains how these results are constructed, firstly explaining how the necessary ranges are derived, secondly explaining how the projections are constructed and thirdly how the results are cautiously interpretable in terms of CO_2 emissions.

2.7.1 DETERMINING RANGES

The probabilistic projections are constructed by combining many model runs in order to create a measure of error on output variables, given an error on the input variable. The input variables used with ranges are the independent variables that are not calibrated in each model run, a listing of which can be found in section 0. The errors on the input variables are henceforth called ranges. These ranges are generated using a combination of sources that have to be congruent for a range to be considered valid:

- The literature values defined in an earlier section, spanning many authors (New & Hulme, 2000)
- The values provided by the authors
- A good historic fit of the model
    - Optimized values generated by the model
    - Sensitivity analysis on the error using Monte Carlo methods
    - Investigation of regional variation if necessary

The literature and author provided values with ranges provide an outer maximum to the ranges. The analysis using a good historic fit of the model are largely based on a publication by Van Ruijven (Ruijven et al., 2009). The quasi-Monte Carlo approach using both random variables and a CCC design spanning the extremes of the variable space has been replicated from van Ruijven. This approach is computationally efficient and the outer extremes of any given simulation are no longer stochastic.

The models used in the Van Ruijven investigation have a very rough variable landscape, with many local minima the optimization algorithm can end up in, instead of simply having a global minimum. Each local minimum is optimal for the set of variables given but not a global solution. By investigating the landscape in this manner an indication of where the global minimum might reside is generated in a relatively computationally efficient manner as well as a ranges which can be weighted depending on the density of solutions.

Although the methodology used in this paper is similar to the aforementioned method by Van Ruijven, the models discussed here turn out to have a smoother variable space and because computational power, due to the supercomputer used in some model runs, is no issue a global solution can be calculated accurately using a genetic algorithm. They can hence be optimized fully for a large number of data points, unlike the Van Ruijven publication where limitations due to the model and computational power reduced the number of model runs to the order of one

hundred. Trail runs with the method for each of the models ended similarly, as depicted in Figure 28 with a probabilistic projection, in the same minimum.

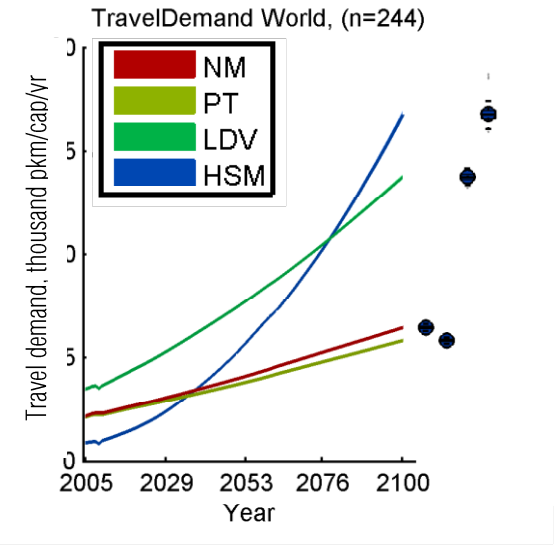


Figure 28: Broad range Monte Carlo simulation over all 7 independent variable, the values are optimized every model run, producing a probabilistic projection as used by B. van Ruijven. It can be seen that optimization gives the same result each time. GCAM

It is therefore decided to modify the method toward more of a classic sensitivity analysis (Hamby, 1994). The methods uses a Monte Carlo simulation over a single dimension for a single region, keeping other dimensions the same and allowing for 35 years of calibration for the calibration variables. The variables are also analyzed across regions, by optimizing this single variable using the methods mentioned in the calibration section. Values for the least uncertain (e.g. ϵ for GCAM) variables are determined first and later to more uncertain variables (e.g. σ).

An example is given in Figure 29 for the ϵ of the GCAM model. The two graphs in the center box of the figure represent an optimization over the value of the independent variable indicated in the x-axis and a modified sensitivity analysis where a Monte Carlo simulation is performed ceteris paribus (i.e. all but 1 variable is modified and the error plotted as a function of this variable).

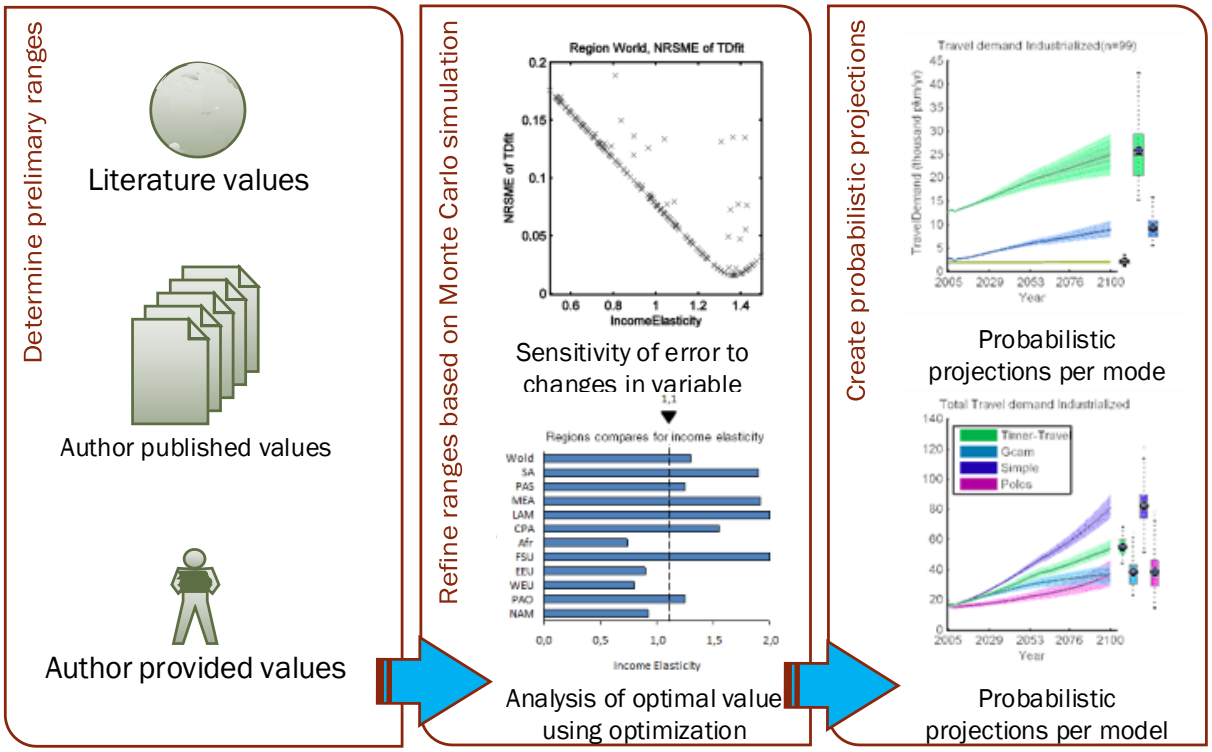


Figure 29: Pipeline for determining ranges, example based on GCAM

2.7.2 CONSTRUCTING PROBABILISTIC PROJECTIONS

In order to construct probabilistic projections the found ranges are combined with quasi-Monte Carlo methods used in earlier investigations, using a CCC (Centered Cubic Structure) design combined with random variables (Ruijven et al., 2009). This combination ensures both a good sampling of the space and all its boundaries, and is therefore superior to purely random Monte Carlo, as described in the previous section.

The ranges are used as values to the input parameters of the models and they calibrate and iterate as normal. The models are run between 100 and 5000 times, depending on the number of variables used in the Monte Carlo simulation. The output is then filtered by eliminating outliers based on either being >95% of the range or more than 1.5 times the distance between the first and third quarter removed from the median solution. This is mostly needed to prevent outliers, caused by either the genetic algorithm or model instability, from affecting the outcome significantly. Additionally only model runs with an error below a threshold, as described in the previous section, are accepted. An example of the design of a probabilistic projection for TD for different modes can be found in Figure 30.

The probabilistic data can be combined across all regions to produce worldwide TD projections similar to earlier publications for the energy-system (Van Vuuren, 2007). The data is multiplied by population in order to get the total TD per region. In case of the developing region an average

value is applied to China (CPA) in order to make the total pkm add up. A threshold is again applied to the data, but slightly looser than before (e.g. 15% for TIMER-travel).

Combining probabilistic data is not trivial, because there are statistical considerations. One could use standard deviation weighing, multiplication per quartile, or other statistic methods (Moss et al., 2010a). In this investigation a simpler approach is used, the data is simply added up per model run and a spread calculated based on the resulting data points. An example of the combined result can be seen in Figure 31.

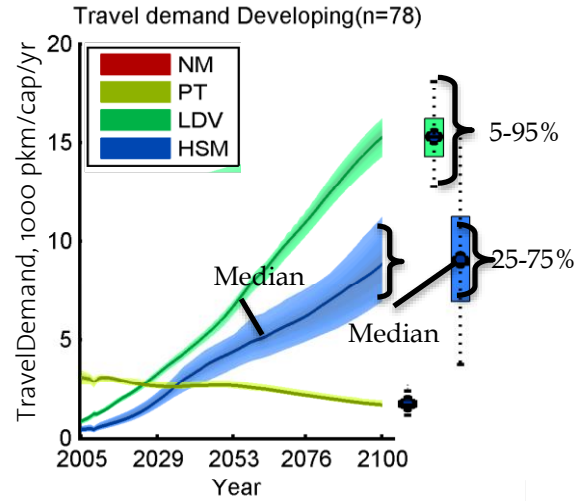


Figure 30: Design of the probabilistic projections, the colored lines represents the modes, indicated percentages are the first and second quartiles. The plot presented is from the Timer-travel model with convergence of preferences. The number of runs in the simulation is given by n.

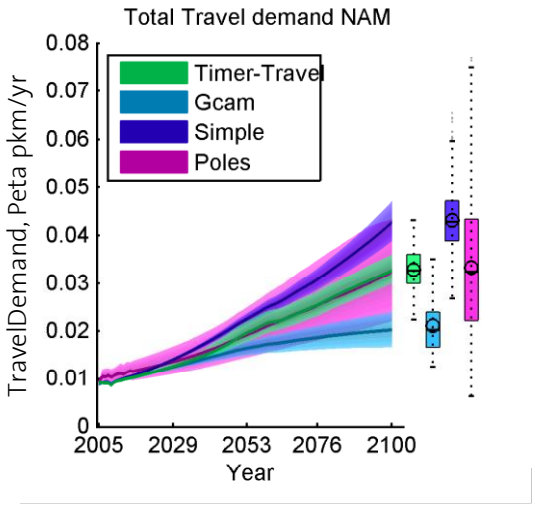


Figure 31: Example of combined probabilistic projections for a single region. All the models are plotted and the travel demand is summed over the population of the region.

The projections can be combined into economic regions, Industrialized, Reforming, Developing and Worldwide, as depicted in Figure 84. Notice that because there is no data for Central Planned Asia the average pkm/yr for the remaining developing nations has been used to fill in for the population gap hence created. Also, for the other aggregated regions not all models are available due to the error rejection process.

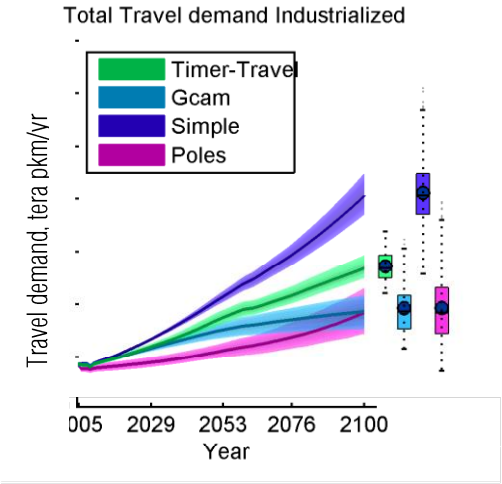


Figure 32: Example of a probabilistic projection for TD with combined regions

3 RESULTS

The results section will first describe the calibration and validation of results for each of the models, including sensitivity analysis and comparison to published values. Figure 33 displays the structure of this section and relates it to the central research question.

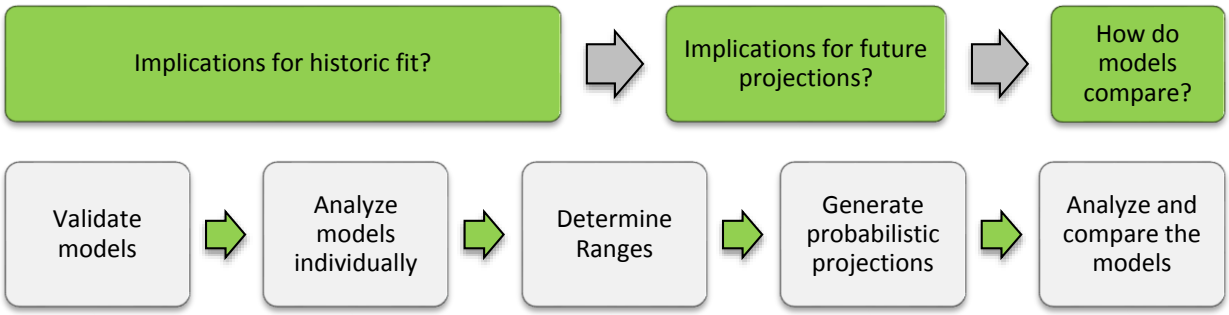


Figure 33: Flow chart of the general methods of this investigation. The top flow denotes the research questions, the lower one denotes the results

1) Validation

   This section will present the results of the validation of each model. Each model depicted here is based on the description in section 2.1.1. Additionally the datasets used and the literature values for the ranges of independent variables per model can be found in section 0. This section will first, per model, present the historic fits, by presenting historic graphs with predicted and actual TD and by presenting the quantitative error (NRSME). Secondly a comparison to published results, if available, will be presented. Lastly the sensitivity analysis will be presented per independent variable, per model.

2) Analysis of individual models

   Before comparing the models, first each of the models is described individually. The results of thesis individual analysis will later be used to compare the models. Not all analyses are interesting for all models, hence many of the performed analysis are omitted in this presentation.

3) Determination of ranges

   This section presents the derivation of the ranges for the probabilistic projections. The procedure used to derive this is presented in section 0. Notice that not all analyses conducted are presented here, only the ones deemed insightful. Regional variation and probabilistic projections are therefore presented only for select variables of interest.

4) Probabilistic projections

   This section presents the probabilistic projections. A combination of the three previous sections and the main result of the present investigation. The TDfits for each model and region can be found in the preceding section. The regions shown here are the world region (a combination of all regions except CPA) or the industrialized region (for POLES and SIMPLE,

as they do not have all regions), as well as three individual regions, Western Europe, Eastern Europe and Sub-Saharan Africa. The latter three regions are chosen to represent the industrialized, reforming and developing region, respectively, and the latter two are some of the tougher fitting regions and hence the reader is presented with a transparent overview. The regions are plotted even if the prior section determines they're unsuited due to a poor fit with known historic facts, for consistency.

- Comparison of the results for each model
    - The results of the individual analyses of the models are compares, these do not include probabilistic projections
    - The probabilistic projections of the original models are presented and compared
    - The probabilistic projections of the models with convergence to industrialized preferences are presented and compares

3.1 THE SIMPLE MODEL

3.1.1 VALIDATION

The SIMPLE model is calibrated by allowing the elasticities to be optimized over the range provided by the literature study (section 2.2.1). This means the outcome of each mode is forced invariant of the historic datasets, therefore some flexibility had to be added. Results are plotted in Figure 34. The error, TDfit, for Light Duty Vehicles LDV and High Speed Modes HSM is very good. It can be seen that the assumption of a negative income elasticity ϵ for the public transport modes, as frequently described in literature, e.g. (Litman, 2011, 2012c; Paulley et al., 2006), might not be warranted for all regions, this will further be discussed in the discussion section.

The assumption of a negative ϵ for Public Transport PT causes a significant error in the fit. Kitous (2000) describes the ϵ as being positive. Schafer takes a central position, travel demand TD for public transport PT initially increasing with income I ($/cap/yr) with and then decreasing (Schafer, Heywood, Jacoby, 2010). This discussion will be continued in the discussion section, for now it is therefore decided to make this ϵ positive and calibrate it based on the average for goods in general (≈1) (Kyle & Kim, 2011).

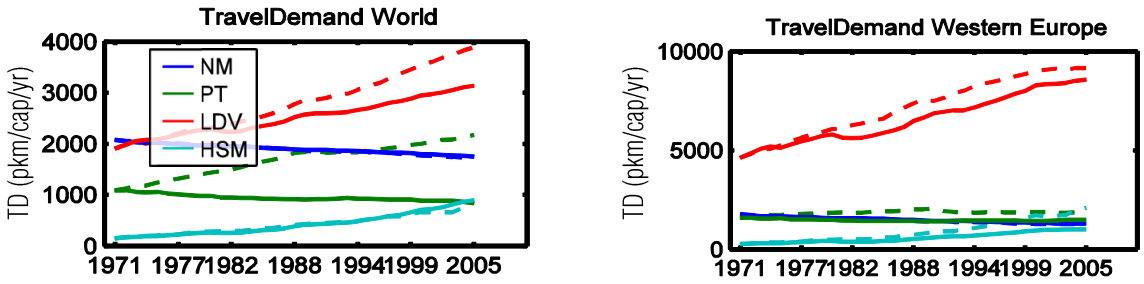


Figure 34: Historic travel demand based on calibration from 1971 to 1980 and projection from 1981 to 2005. The dotted lines are historic data. From left to right the errors, TDfit, are 10%, and 10%. SIMPLE

The result of correcting the income elasticity for PT can be seen in Figure 35. The error, TDfit, is much more acceptable in this setup, for Africa a fit of the PT mode is far from perfect. This is likely due a low value for income elasticity in Africa.

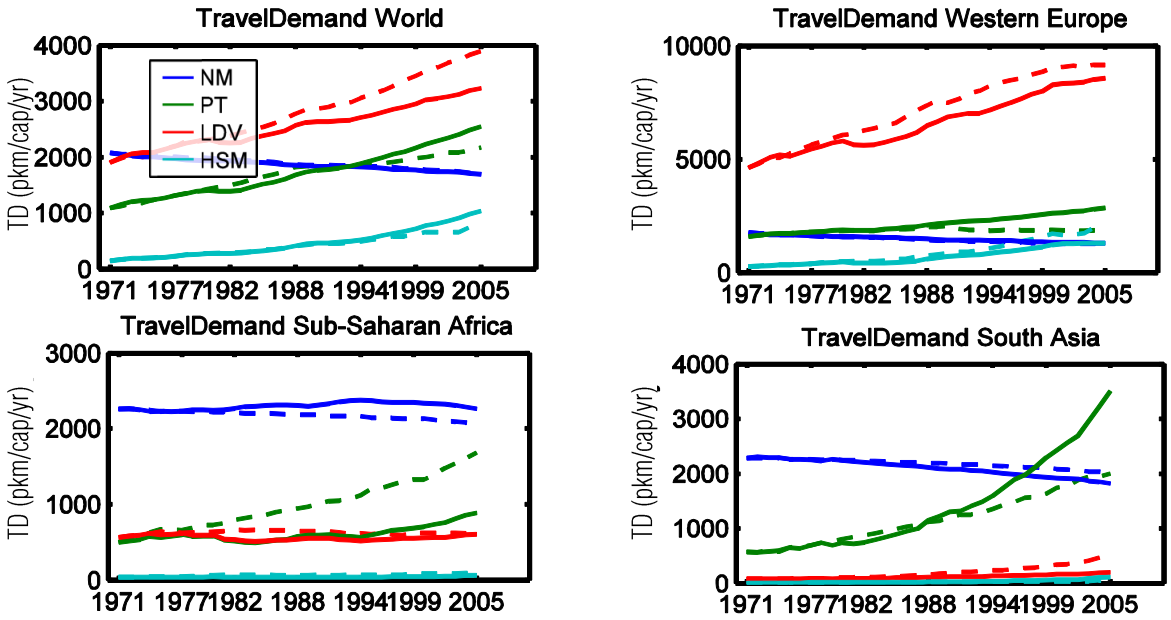


Figure 35: Historic travel demand based on calibration from 1971 to 1980 and projection from 1981 to 2005. The dotted lines are historic data. From left to right the errors, TDfit, are from left to right 3.4%, 4.8%, 9.9% and 7.3%. ϵ for PT is taken positive. SIMPLE

Figure 36 contains the errors for the Monte Carlo simulation based on the ranges given earlier. It can be seen that, the results fit with the historic data quite well. Remarkable considering the simplicity of the model. In fact for the industrialized and reforming region the values are within bounds (<10%) to be considered in the global assessment. The developing region almost all have a TDfit above the threshold and is therefore decided to reject all the developing region.

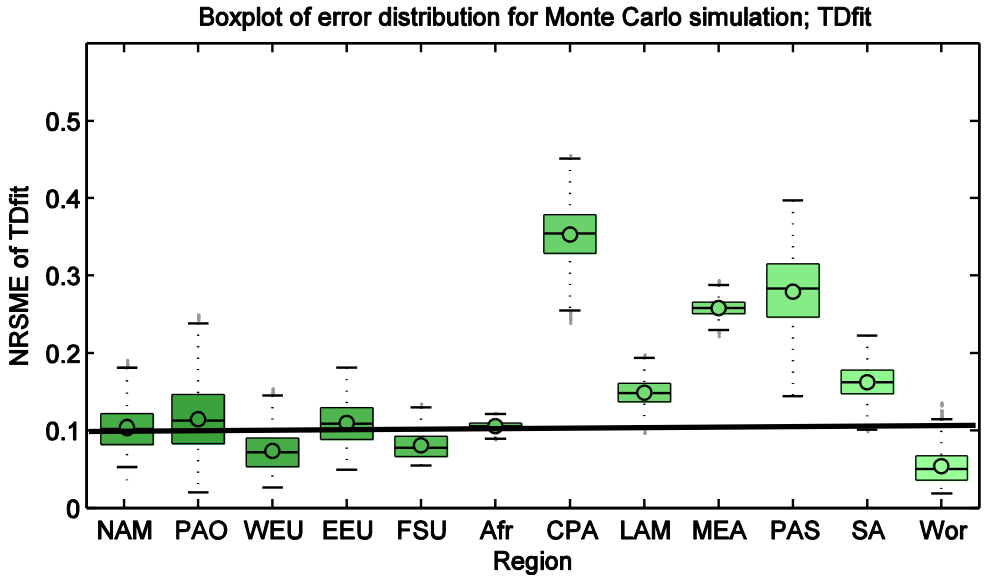


Figure 36: Boxplot of the TDfits over the unfiltered range used in the Monte Carlo analysis per region. The line indicates the threshold for filtering. SIMPLE

3.1.2 ANALYSIS

The sensitivity of the travel demand in 2100 is next analyzed. First it can be seen in Figure 37 that the sensitivity to the income elasticities ϵ is high, except for the non-motorized NM, this is probably due to its small size in the world region. Figure 38 shows the sensitivities to travel price elasticities σ, it can be seen that especially LDV is sensitive, this is due to high TD for LDV in 2005.

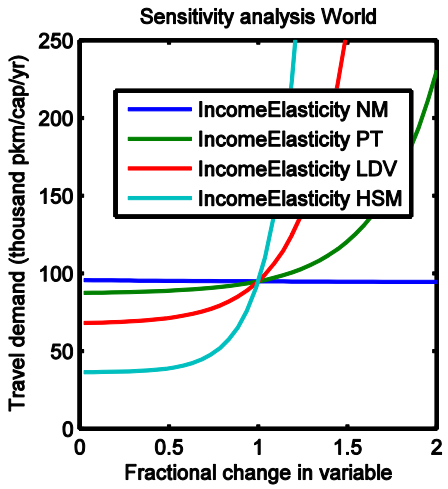


Figure 37: Sensitivity analysis for ϵ. Non-normalized values are -0.2, 1, 1.1, 2 for NM, PT, LDV and HSM, respectively. SIMPLE

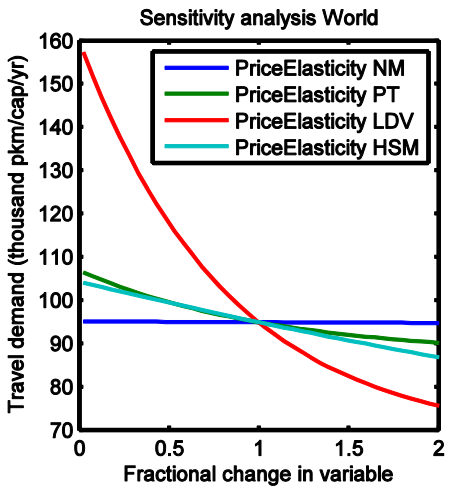


Figure 38: Sensitivity analysis for σ. Non-normalized values are 0.5, 0.4, 0.8, 0.75, for NM, PT, LDV and HSM, respectively. SIMPLE

The Total Travel Cost TTC (% of income I) and Time Use TU (hr/cap/yr) values for the simple model are next analyzed. As before the values from the best fitting runs of the Monte Carlo simulation are used. Figure 39 shows that the bulk of the TTC values are below the maximum

value of the Travel Money Budget TMB (9-15% of I) (Schafer, Heywood, Jacoby, 2010). The Developing region don't adhere to the TMB as strictly, the same regions also have a high error, TDfit. WEU has a high TTC, however these are forecasts that might not be unacceptable in a future scenario. Figure 40 depicts the TU for each of the modes. For all regions but Africa it can be seen that the maximum Travel Time Budget TTB found in literature (1.5hr/day, (Avineri, 2009)) is exceeded. This seems unrealistic but needn't be impossible, traveling 2.5 hours per day, as predicted for WEU is unlikely but not impossible. The model predictions are still produced and presented but should hence be taking with more caution.

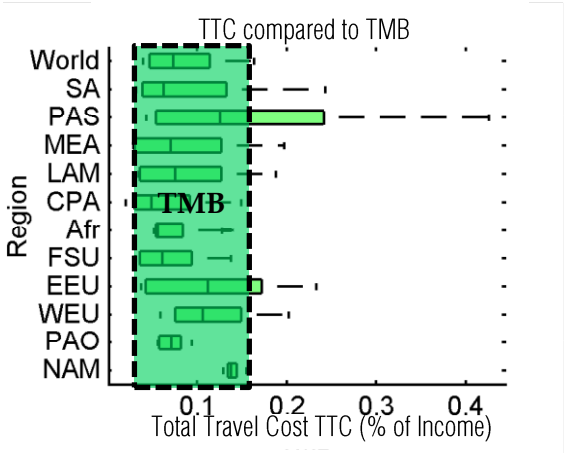


Figure 39: Boxplot of TTC for 130 years for the best Monte Carlo run of all regions. Dashed box represents the range of found values of TMB. SIMPLE

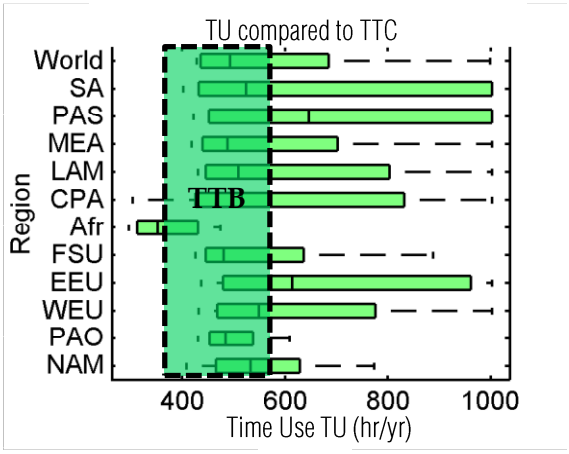


Figure 40: Boxplot of TU for 130 years, motorized modes, for the best Monte Carlo run of all regions. Dashed box represents the range of found values of TTB. SIMPLE

3.1.3 RANGES

The simple model lacks any form of calibration on historic data. The filtering of the Monte Carlo results will therefore be the only check on historic fit. The values are determined using the method described in section 0. Presented are the values produced when optimizing each independent variable to each region ceteris paribus, Figure 41, as before. It can be seen that the optimal values for each variable and mode differ per region.

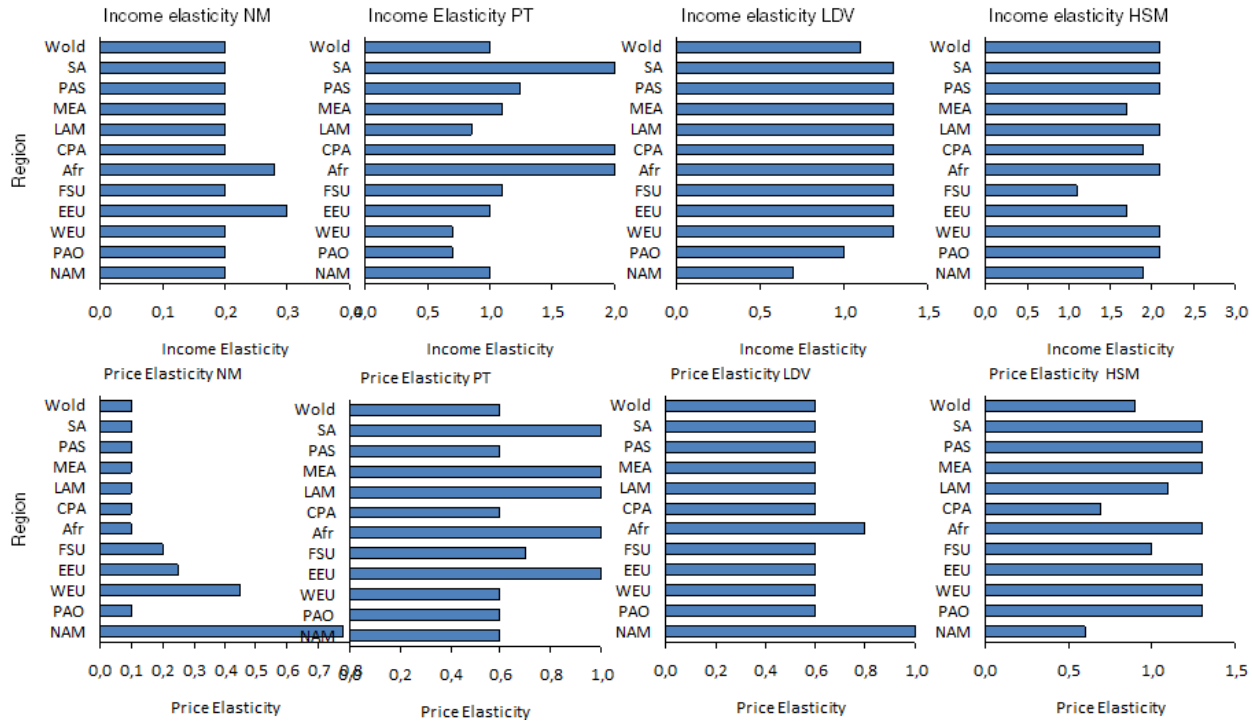


Figure 41: Investigation of the values produced when running optimizations over each variable, ceteris paribus. SIMPLE

The range can be found in Table 13.

Independent variable	Symbol	Value	
Travel Price Elasticity	σ	NM:	-0.3± 0.1
		PT:	-0.7± 0.2
		LDV:	-0.6± 0.2
		HSM:	-1 ± 0.3
Income Elasticity	ϵ	NM:	-0.2 ± 0.1
		PT:	1 ± 0.3
		LDV:	1.1 ± 0.2
		HSM:	2.0 ± 0.3

Table 13: Ranges used for the SIMPLE model

3.1.4 PROBABILISTIC PROJECTIONS

Figure 42 displays the probabilistic projections for the SIMPLE model. It can be seen that the formulation of the model leads to fairly straightforward projections. The results are therefore very intuitive. For all regions the travel demand of LDV, TP and HSM increases and for NM decreases. What is surprising is that for the industrialized region the travel demand for HSM does not overtake the TD for LDV in Western Europe. Also the PT mode remains very significant throughout Western Europe and Africa.

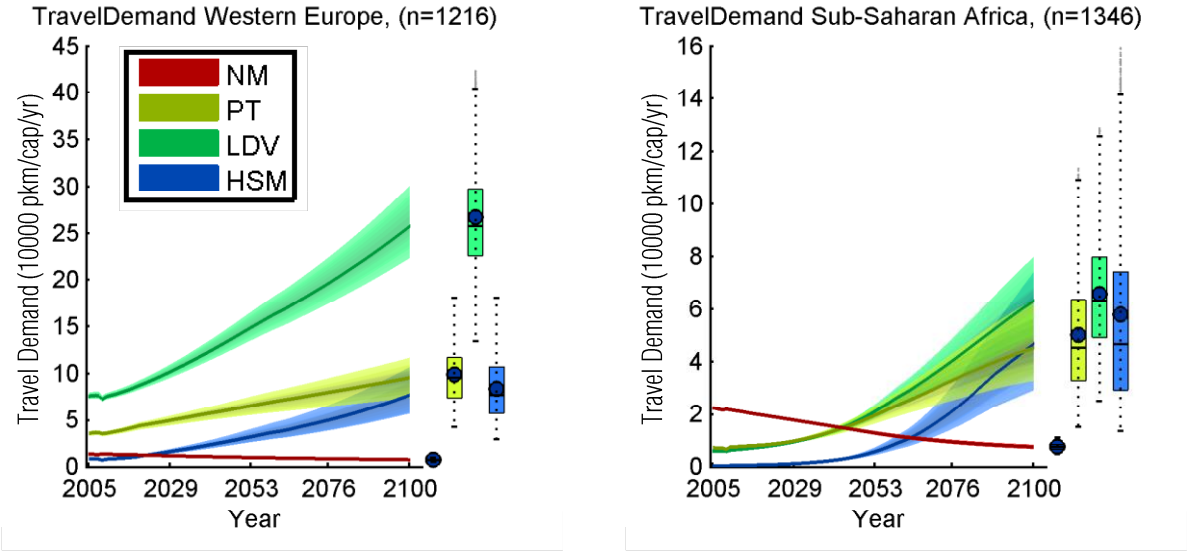


Figure 42: Select probabilistic projections, the error threshold is increased from 10% to 20% for the developing region. SIMPLE

Figure 43 depicts the aggregated projections. Notice that only industrialized and reforming region are considered valid based on historic fits. The travel demand of industrialized region appears to grow strongly for LDV and HSM, whilst keeping PT stable. No saturation can be seen, and HSM should become the largest mode in the 22nd century. Reforming region show greater diversity, with TD for PT of significant size, by 2100 HSM will overtake LDV, and there is no saturation effects.

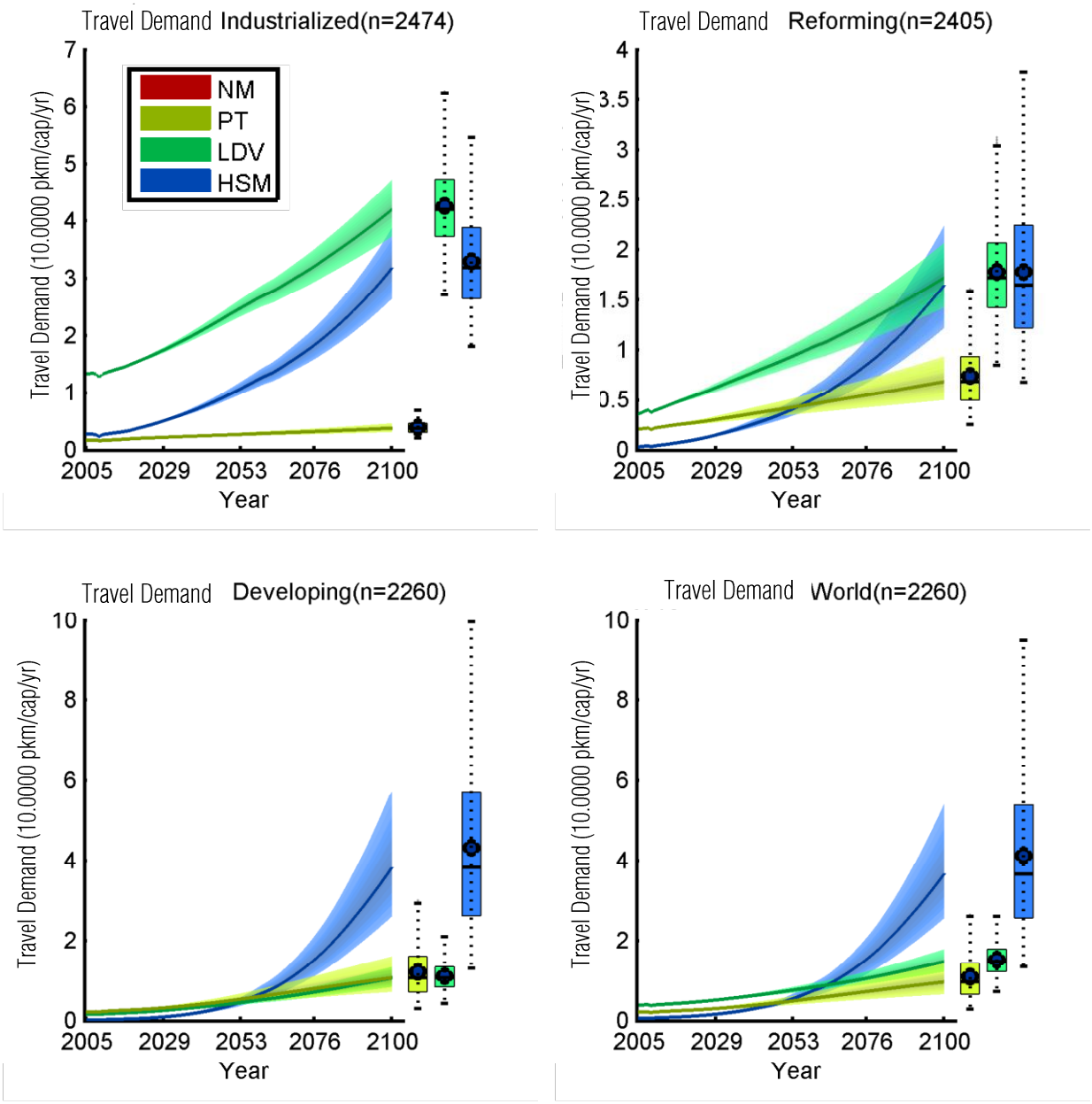


Figure 43: Probabilistic projections for the aggregated regions. SIMPLE

3.2 THE POLES TRAVEL SUBMODEL

3.2.1 VALIDATION

The historic projection for the POLES model is depicted in Figure 44. The time trend TT is calibrated over the years 1971-1980 for all regions. The error, TDfit, is sufficiently small for the four graphs presented. Deviations from the historic set notably appear closer to the final year of the projection, 2005. There is no specific direction, higher or lower, to these changes compared to the historic set. It can be hypothesized that the static value of the elasticities are the cause, this will be described in the discussion. This effect can for example be seen in Africa where the elasticity for HSM appears overestimated for the range 1980-2005. TD for HSM grows much faster than the actual historic set, possibly due to the same reason.

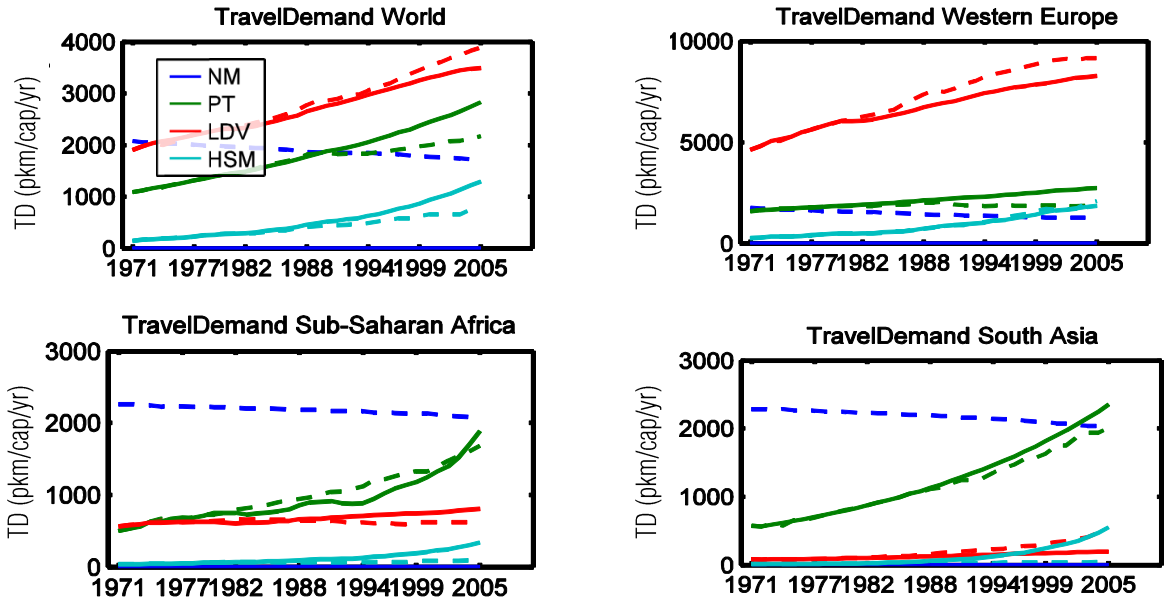


Figure 44: Historic travel demand based on calibration from 1971 to 1980 and projection from 1981 to 2005. The dotted lines are historic data. The errors, TDfit, are from left to right, 5.1%, 2.7%, 8% and 8.2%. POLES

Depicted in Figure 45 are the errors for the POLES model, based on the ranges determined. It can be seen that most of the regions to not fit to the criteria mentioned. Only the Industrialized region produces a good fit. The causes and solutions for this will be discussed extensively in the discussion and conclusion sections of this research. It is decided that only the industrialized region, NAM, PAO and WEU will be used for comparisons using the POLES model.

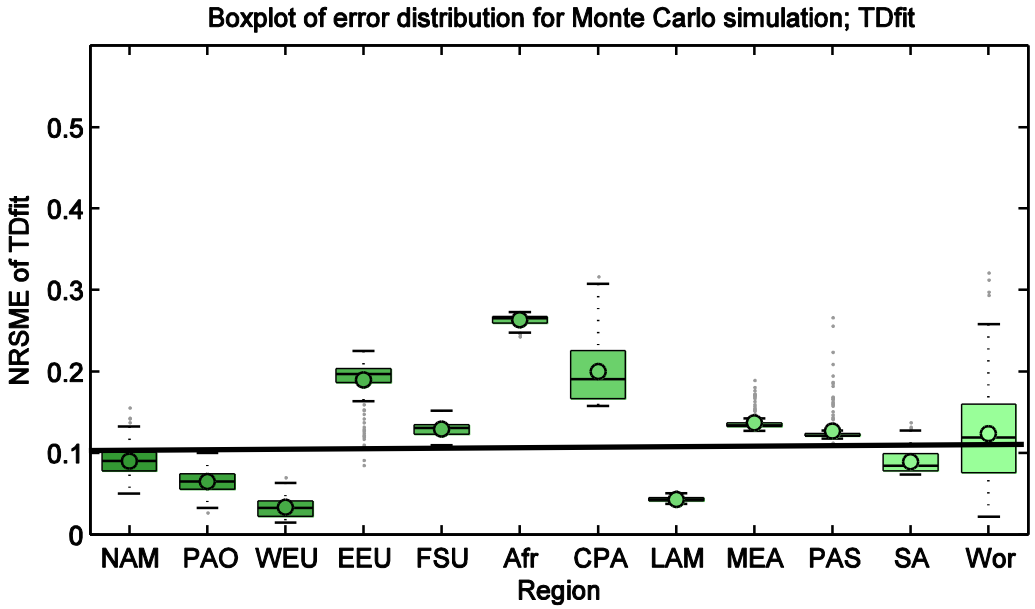


Figure 45: Boxplot of the TDfits over the unfiltered range used in the Monte Carlo analysis per region. The line indicates the threshold for the filtering. POLES

3.2.2 ANALYSIS

The sensitivity analyses are next presented, first Figure 46. It can be seen that TT strongly affects the outcome of the model. TT is here kept negative because a TT of positive 1% for the world region causes values to go toward infinity due to the way the saturation is formulated, see section 3.2.2. In Figure 47 the sensitivity for the saturation of HSM (TDSAT) is high; in fact the increase in travel demand for HSM in this region is so strong that the saturation is the limiting variable. Income elasticity on the equipment rate of cars ϵER is only of effect if its value is small, it doesn't restrict otherwise.

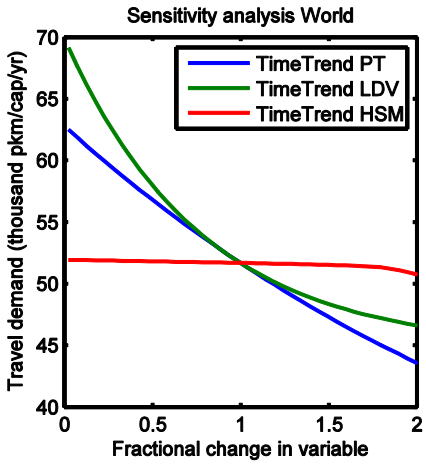


Figure 46: Sensitivity analysis, non-normalized values are -1% for the TT for all modes. POLES

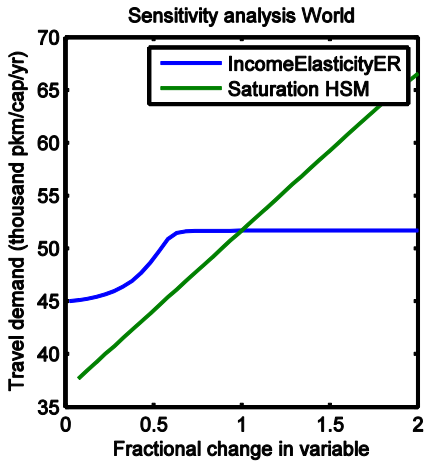


Figure 47: Sensitivity analysis, non-normalized values are 1 and 50000, ϵER for LDV and ERSAT for HSM, respectively. POLES

Figure 48 depicts the sensitivity to σ for each mode, σ for PT is positive and σ for LDV is negative, the TD in 2100 is sensitive to them. TD is only weakly sensitive to σ for HS, this is due to the saturation limit TDSAT. The sensitivity to ϵ is depicted in Figure 49. It can be seen that the model is sensitive to the σ of the PT and weakly sensitive to the σ for LDV and HSM. This is due to the fact that the elasticity on PT is not bound by saturations.

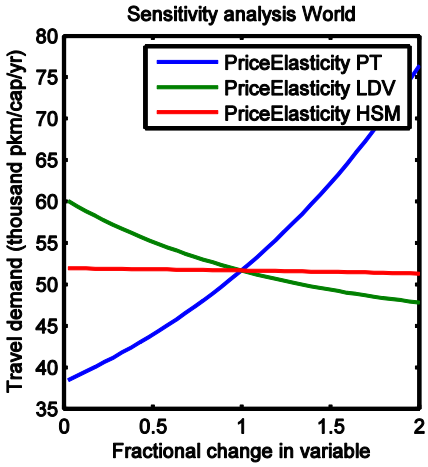


Figure 48: Sensitivity analysis, non-normalized values are 0.2, -0.4 and -0.4, for δ of PT, LDV, HSM, respectively. POLES

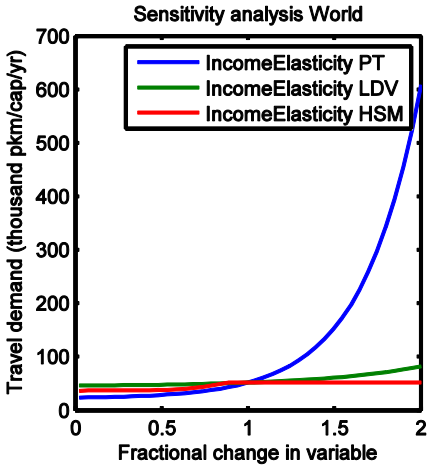


Figure 49: Sensitivity analysis, non-normalized values are 1.0, 1.1, 2.0, for ϵ of PT, LDV and HSM, respectively. POLES

Depicted in Figure 50 is the TTC for all regions' best fitting runs during the Monte Carlo simulations, it can be seen that for many developing region the TU > TMB, TU is larger than Schafer's 9-15% (Schafer, Heywood, Jacoby, 2010). The TU is depicted in Figure 51, it can be seen that for all non-industrial regions TU is higher than the maximum TTB from literature (Avineri, 2009; Susilo & Avineri, 2011).

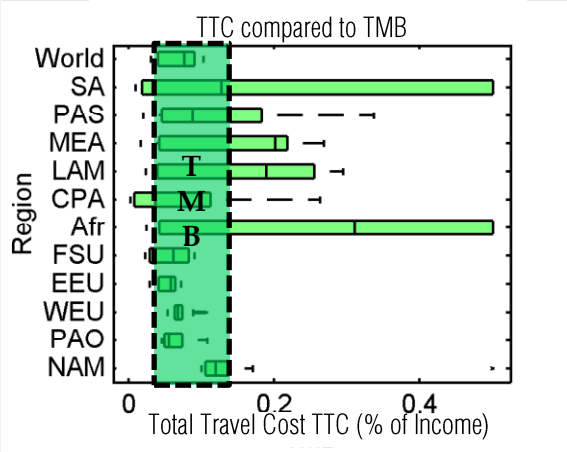


Figure 50: Boxplot of TTC for 130 years for the best Monte Carlo run of all regions. Dashed box represents the range of found values of TMB. POLES

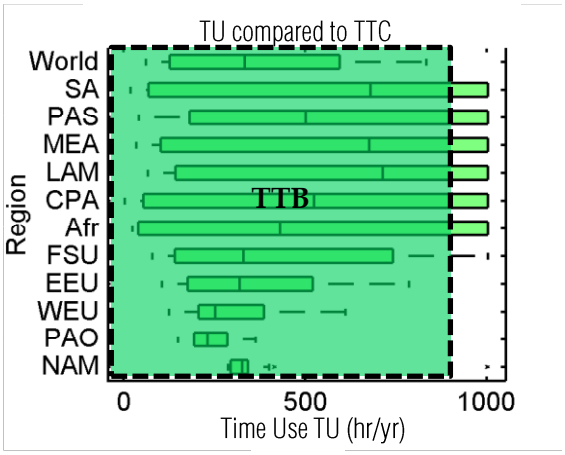


Figure 51: Boxplot of TU for 130 years, motorized modes, for the best Monte Carlo run of all regions. Dashed box represents the range of found values of TTB. POLES

In order to restrict the individual modes from reaching unrealistic values the POLES model employs a limit on the number of cars that are available in a region. The equipment rate ER therefore tends to develop once the income has reached approximately 5000$/capita and moves to a saturation limits of some 0.9 vehicles/capita as income develops (Chamon, M. Mauro, P. Okawa, 2008). The exact derivation of this relationship is displayed in the methods section.

There is literature that describes both the development of the ER and the saturation limit ERSAT in detail, as depicted in Figure 53. The POLES model also assumes a development, as depicted in Figure 52. It becomes clear that the derivation of ER by the POLES model only matches that of detailed models in other literature (Chamon, M. Mauro, P. Okawa, 2008) for the industrialized region, the regions with the highest equipment rate.

The relationship for the ER, the TU and TTC analysis and the error analysis done earlier all provides the same conclusion. Only industrialized region abide historic findings and the empirical facts of ER, TTB and TMB.

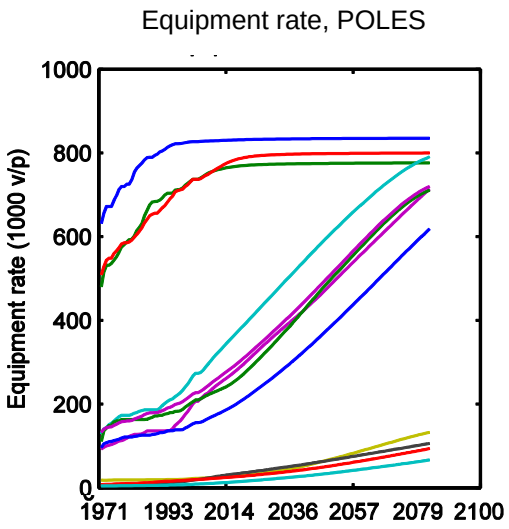


Figure 52: ER for LDV as modeled by the POLES model (Kitous, 2010a)

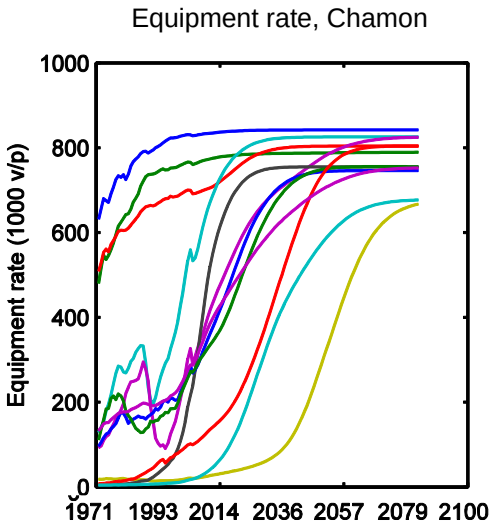


Figure 53: ER for LDV as found in an econometric investigation (Chamon, M. Mauro, P. Okawa, 2008)

Two saturation limits exist in this formulation of the POLES model for LDV and HSM. Shown in Figure 54 is the behavior of HSM when the TT is positive 0.5% and increases the size of the HSM mode over the saturation limit TDSAT, this causes TD to increase to near infinity just over the threshold. This behavior causes instabilities in the model if the TT is not restricted.

Depicted in Figure 55 is the development of TD (for NAM), it is clear that once the saturation ERSAT is reached ER will no longer increase, but because ER and TD are coupled, increasing price will lower the travel demand due to the fuel price elasticity δ. Literature by Schafer expects a similar progression for LDV (Schafer, Heywood, Jacoby, 2010), but not this early on in history.

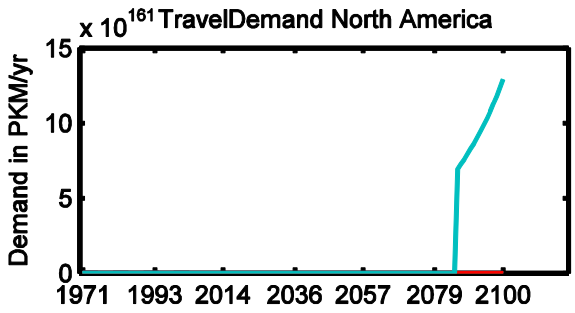


Figure 54: Effect of exceeding TDSAT on the development of HSM. POLES

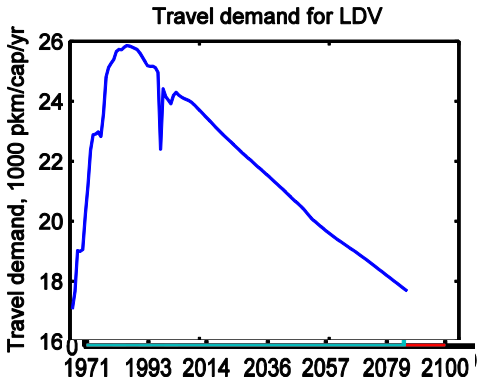


Figure 55: Effect of the ERSAT for LDV on total travel demand (region NAM). POLES

3.2.3 RANGES

In determining the set of ranges for the Monte Carlo analysis, three independent variables are investigated. The income elasticity ϵ, income elasticity on ER ERϵ and the fuel price elasticity δ.

ϵ for PT is positive here. It should be negative based on all literature sources consulted, notably (Litman, 2011; Paulley et al., 2006). However, model documentation clearly states this should be positive, therefore a value for ϵ (Schafer & Victor, 2000) of 1.0 is used instead, similar to the value in SIMPLE. The validity of this across regions is investigated in Figure 58. It is clear that the value varies greatly across regions, and can be both negative and positive. It is clear that for many regions this positive elasticity does not hold.

It is clear that all modes are sensitive to the ϵ, but not so sensitive as to warrant large deviation from literature value, in brackets. The new values are hence 1.0 ± 0.2 (1.0 ± 0.2), and 1.1 ± 0.2 (1.0 ± 0.3) and (2.0 ± 0.5), for PT, LDV and HSM respectively.

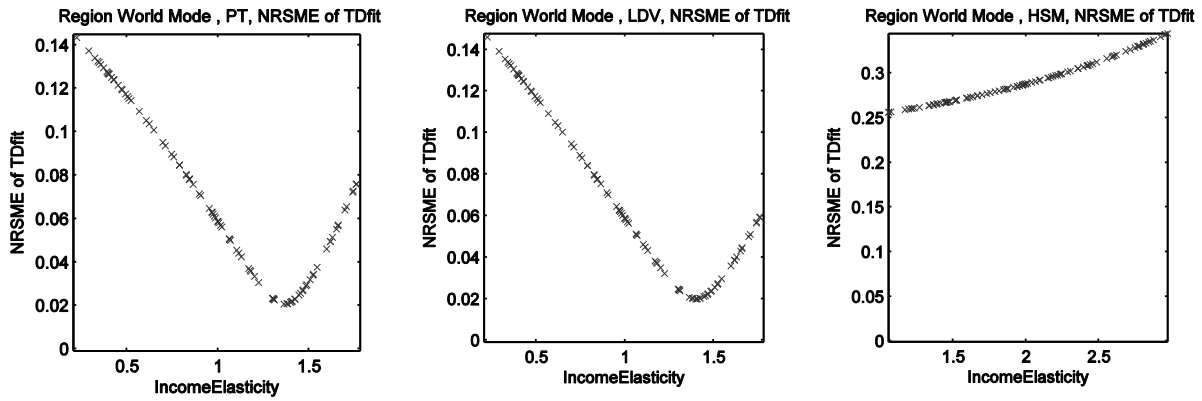


Figure 56: Sensitivity of the error, TDfit, on ϵ for each mode. POLES

Knowing these income elasticities δ is next investigated in an identical experiment to earlier. It can be seen that the fuel price elasticities are insensitive, the error never exceeds 6%. The values from literature are therefore used directly; -0.2 ± 0.1, -0.4 ± 0.2 and -0.4 ± 0.2 for PT, LDV and HSM, respectively.

Depicted in Figure 56 are sensitivities done for the car equipment rate elasticity.

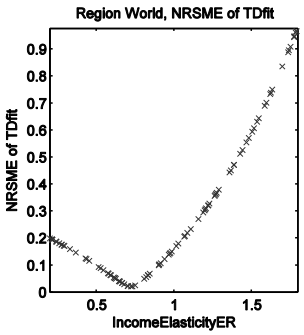


Figure 57: Sensitivity of the error, TDfit, on ERϵ. POLES

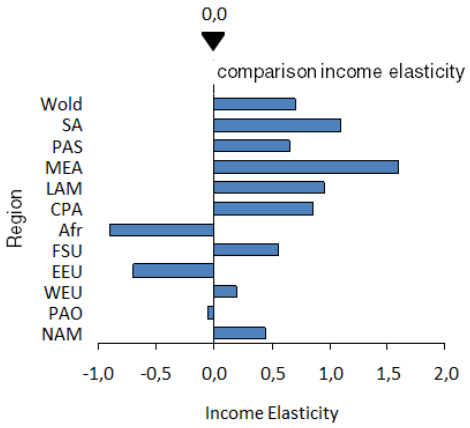


Figure 58: Values of ϵ when optimizing per region ceteris paribus. POLES

The ϵ on the equipment rate is chosen at 1.0 ± 0.2 (literature 1.0 ± 0.25), when the outcome turns out to be very sensitive. This elasticity operates on the income growth and leads to a growth in the ER for LDV. The final ranges for POLES are described in Table 14.

Independant variable	Sym.	Mode	Range
Fuel Price Elasticity	δ	PT	0.2±0.1
		LDV	-0.4± 0.2
		HSM	-0.4 ± 0.2
Income Elasticity	ϵ	PT	1.0 ± 0.2
		LDV	1.1 ± 0.2
		HSM	2.0 ± 0.55
Income Elasticity on Equipment Rate	ERϵ	LDV	1 ± 0.2

Table 14: Ranges, as determined for the POLES model

3.2.4 PROBABILISTIC PROJECTIONS

The POLES model is only valid for the industrialized nations, NAM, WEU and POA. The other regions are not considered sufficiently realistic in reproducing the historic dataset, or empirical findings, they are presented here regardless. As can be seen in Figure 59 Western Europe has projections with increasing travel demand for PT and HSM, with a declining travel demand for LDV likely due to increases in the fuel price. HSM is projected to be the major mode of transportation by 2060, although the spread is very large. Africa produces a high TD for LDV. In this model run inhabitants of the region would, on average have to travel 82 hours per day by 2100 which is nonsense.

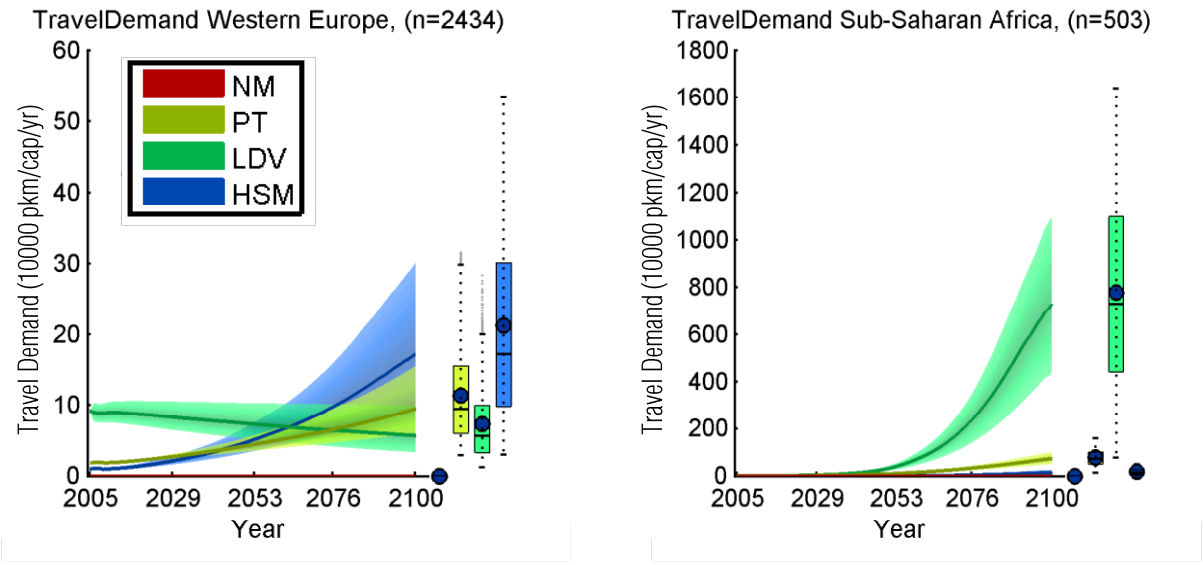


Figure 59: Select probabilistic projections, the error threshold is increased for the developing- and reforming region. POLES

Figure 60 presents the projections for the aggregated regions. It can be seen that for industrialized the TD for HSM is the largest by 2060, with no saturation near 2100. Additionally the TD for PT will overtake LDV, which declines over time. The TD for LDV in the reforming region first increases and then decreases once ERSAT has been reached. TD for PT is the largest mode by the end of the century with a large spread, and especially fat tails (the 95%interval ranges from 10 – 90 Tera pkm). TD for PT and HSM increase at a roughly equal rate, and end up with a HSM roughly 2.5 times smaller than PT and LDV. Additionally, all modes in the Developing region appear saturated by 2100.

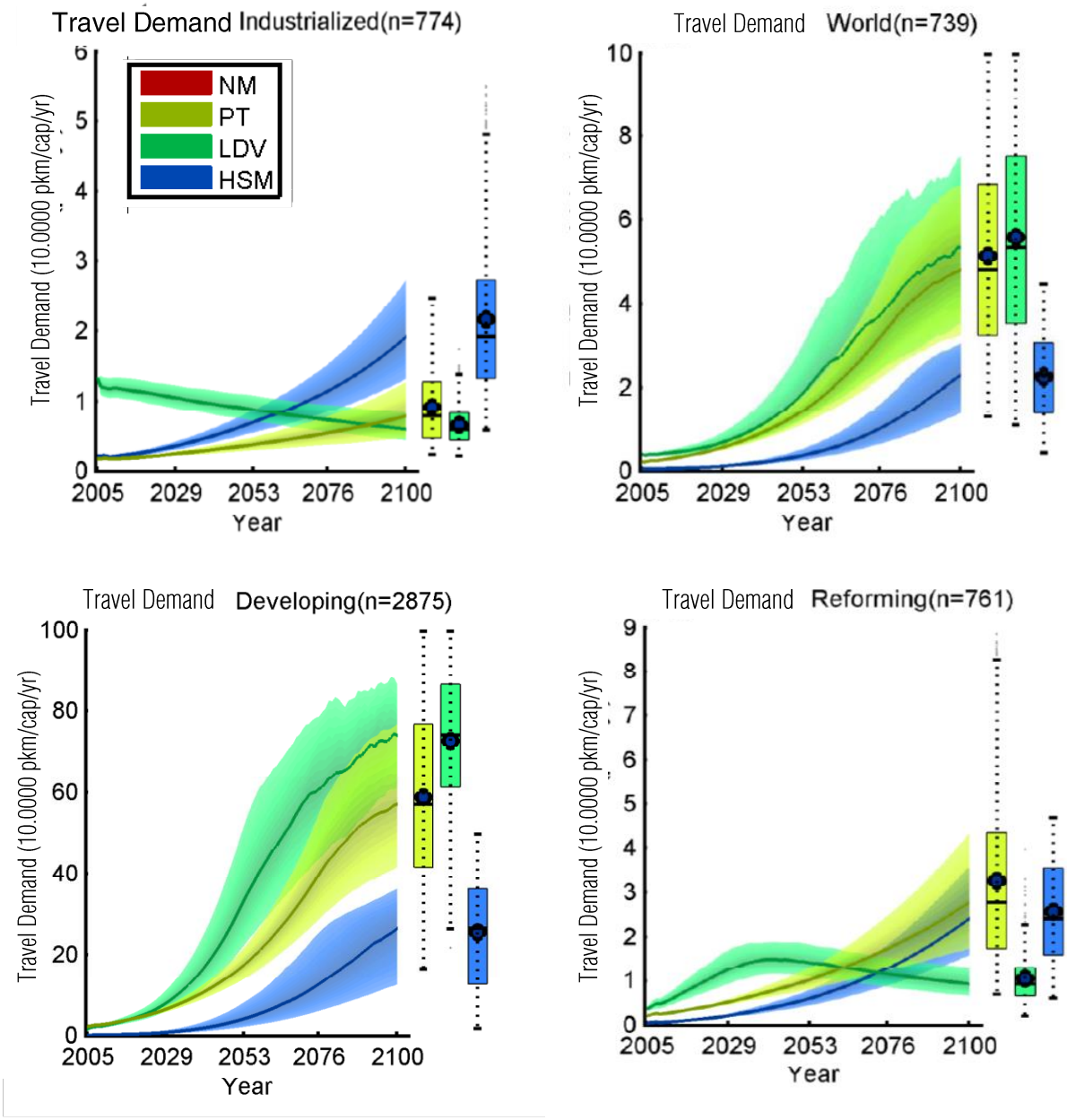


Figure 60: Probabilistic projections for the aggregated regions. POLES

3.3 THE GCAM TRAVEL SUBMODEL

3.3.1 VALIDATION

Non-motorized modes NM have been added to the original formulation (Kim, Edmonds, Lurz, Smith, & Wise, 2006). This was found beneficial to the fit of the model on developing region, and improves comparability to the other models and the TTB, see section 2.2.3. The results without NM can be seen in the appendix, they are almost identical the the model runs with NM

The optimization procedure is changed compared to the procedure published by the authors (Kim et al., 2006) and the general procedure described earlier, as can be read in the calibration section. The consequence is a better historic fit. The historic fit of the GCAM model can be seen in Figure 61. The elasticities were optimized within the literature range to incorporate the slight regional differences in their values. The quantitative fits are sufficient for all but for CPA. Africa has a deviation from the optimal fit for PT, this is due to more rapid growth predicted based on income. The same argument holds for South Asia.

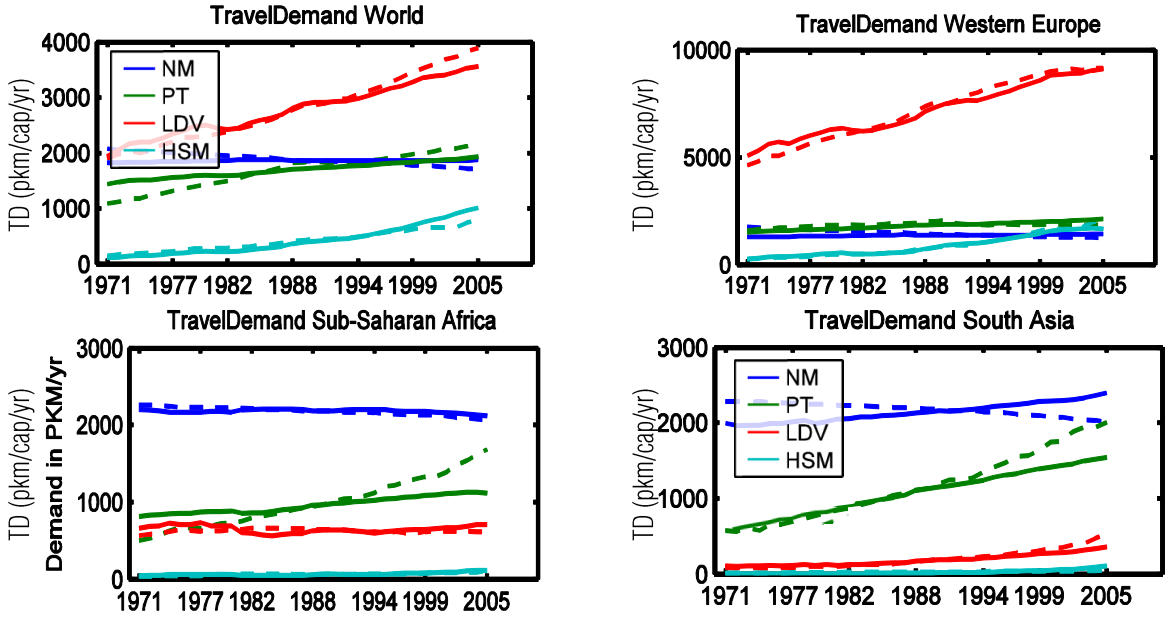


Figure 61: Historic travel demand based on calibration from 1971 to 1980 and projection from 1981 to 2005. The dotted lines are historic data. From left to right the errors, TDfit, are, from left to right, 2.2%, 2.3%, 6.1% and 2.4%. GCAM

In order to ensure the model has been replicated and calibrated properly the results from Kyle & Kim (2011) were replicated. The replication concerns the projection. Therefore a visual inspection is done on several hundred plots in the range provided below. The projections can be seen for the world region in Figure 62, and Figure 63. Other aggregated regions can be found in the appendix. The fit on the separate world-regions (see appendix) gives qualitatively similar results for industrialized and developing region. Including economic crisis in the model accounts for the slightly shifted starting positions of industrialized region.

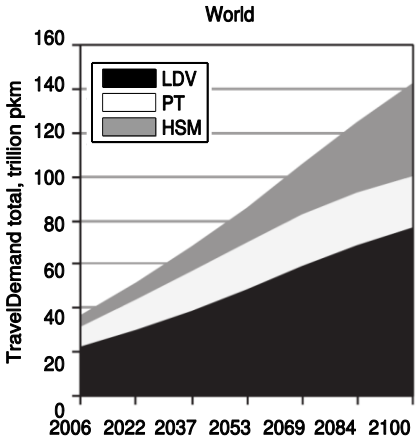


Figure 62: Published results. GCAM

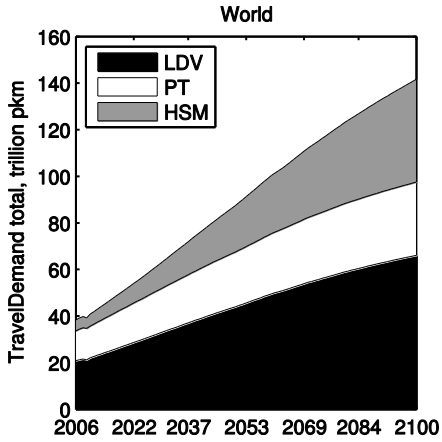


Figure 63: Replicated results. GCAM

The errors on the TD are plotted in Figure 64. It can be seen that except for MEA, all regions have data points below the 10% error margin. MEA is, however, behaviorally sound and therefore not excluded only on the basis of its error. CPA on the other hand has a very large spread, and is behaviorally not sound. It can therefore not be included in the analysis.

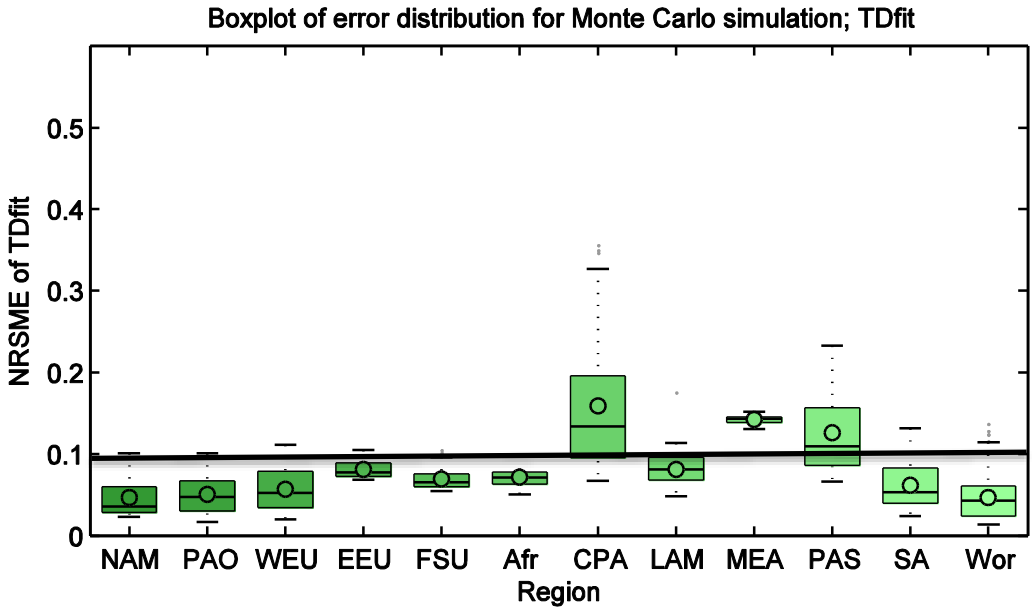


Figure 64: Boxplot of the TDfits over the unfiltered range used in the Monte Carlo analysis per region. The line indicates the threshold for the filtering. GCAM

3.3.2 ANALYSIS

The sensitivity analysis, depicted in Figure 65 and Figure 66, shows that the travel demand is sensitive to changes in ϵ and to some degree to the share weight SW of LDV and HSM. TD is insensitive to σ and the cost distribution parameter μ. The sensitivity of TD to ϵ is of the most significant because of its asymmetry, the distribution is skewed toward the higher values. This

will produce probabilistic projections that have maximum values of travel demand that are significantly further from the median than the minimal values of travel demand. Therefore the maximum values of projections made with GCAM need to be interpreted with care.

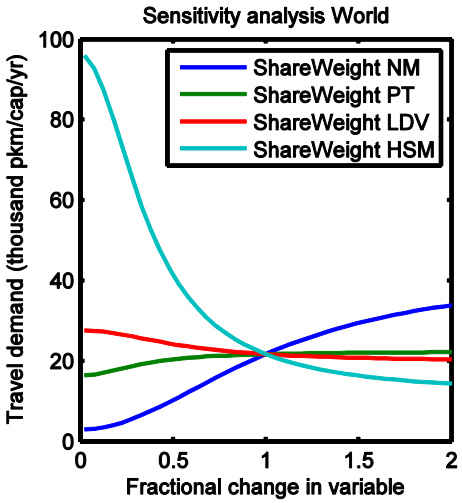


Figure 65: Sensitivity analysis, non-normalized values are 4, 20, 17 and 47 for the SW for PT, LDV and HSM, respectively. GCAM

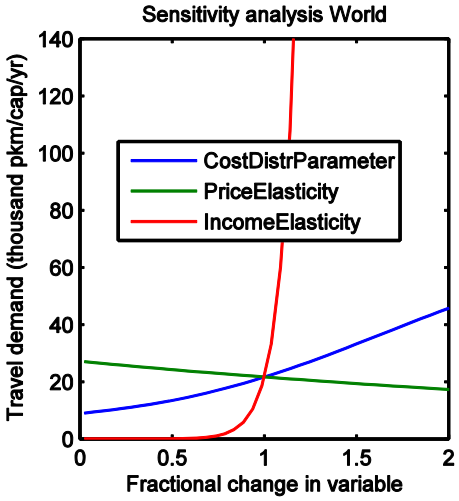


Figure 66: Sensitivity analysis, non-normalized values are -2, 1 and 1 for theμ, σ and ϵ, respectively. GCAM

The GCAM model, as mentioned, does not restrict either Total Travel Cost TTC or Time Use TU. There is feedback through the Cost of Time COT ($/pkm), defined as a function of wage rate WR and DDS, this form is feedback is, however, indirect. Nonetheless the authors of the GCAM model state in their publication that none of their recorded TU values passes the TTB of 1.5 hours per day (~550 hr/yr) (Kyle & Kim, 2011). It can be seen in Figure 68 that some of the data does overshoot this threshold, but the bulk of the data adheres to this restriction. The TTC is displayed in Figure 67, and shows some high values for SA and PAS (Schafer, Heywood, Jacoby, 2010).

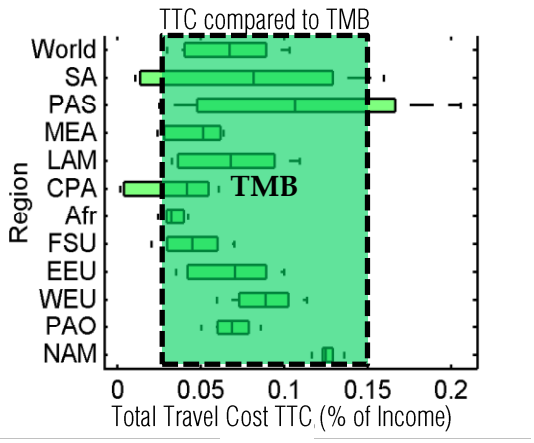


Figure 67: Boxplot of TTC for 130 years for the best Monte Carlo run of all regions. Dashed box represents the range of found values of TMB.GCAM

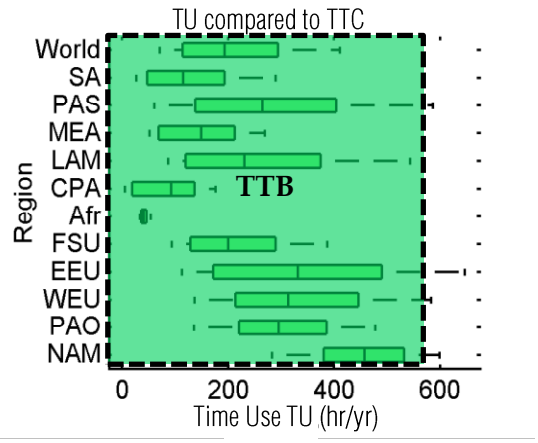


Figure 68: Boxplot of TU for 130 years, motorized modes, for the best Monte Carlo run of all regions. Dashed box represents the range of found values of TTB. GCAM

Closer investigation of SW is presented in Figure 69. It can be seen that the NM have the lowest SW in all regions and PT varies compared to LDV. HSM on the other hand has a high SW for the developing region but is much smaller than the industrialized, the inverse is true for LDV. This mean the preferences differ between industrialized and developing region.

Figure 70 displays the error distribution if the values provided by the authors were used directly. These values are μ = -2, an ϵ = 1 and σ = 1. It can be seen that for the industrialized and reforming region this is a good assumption, but for the developing region this is not always the case. It is therefore crucial to filter the best fits from the ranges provided to ensure a good historic fit. In addition to this it can be seen that the best fit by is for North America, considering the model was probably originally developed for this region this is no surprise.

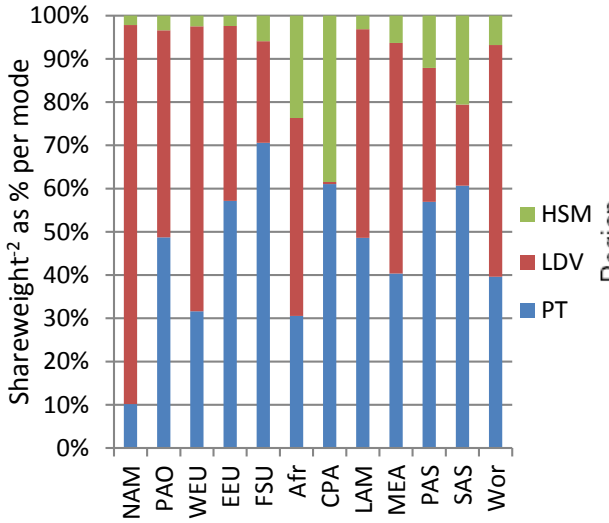


Figure 69: SW for a typical model run, SW is to the power of -2, as in the formulation. GCAM

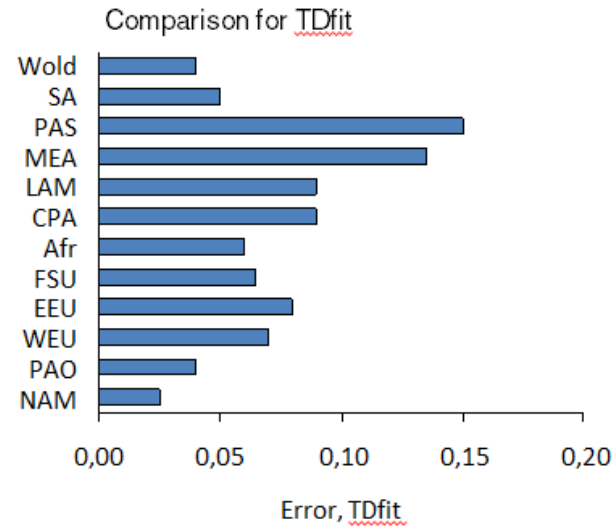


Figure 70: Errors, TDfit, when using the model with author provided values. GCAM

The mode cost MC ($/pkm) for LDV over time is investigated in Figure 71. The prices increase over time, but do not converge to the same value for all regions. The travel price TP ($/pkm) in 2005 for LDV in NAM is in the order of 20 cents/pkm, the MC is 40 cents/pkm. Therefore roughly half of the MC is based on the income effect present in the COT. The growth is therefore largely due to the increase in income and hence WR. Figure 72 depicts the MC for HSM, here the COT is much lower than in LDV, because the transportation is some six times faster. With this component much smaller from the onset the prices do not increase as much as for the LDV mode, at least in an absolute sense.

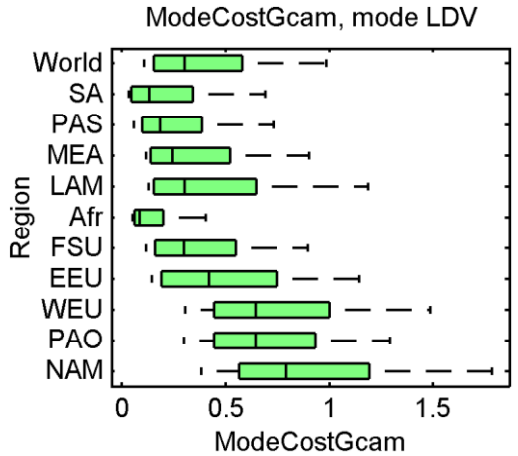


Figure 71: MC for LDV for all regions over time. GCAM

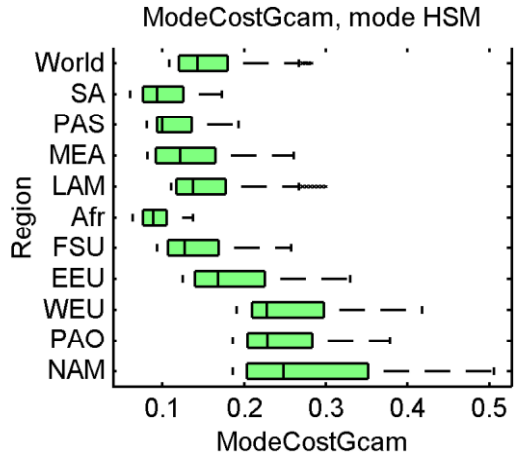


Figure 72: MC for HSM for all regions over time. GCAM

The last investigation concerns the use of increasing door to door speed DDS in the GCAM model. As mentioned the authors assume a constant DDS. The difference between urban- and rural DDS, which tend to change the speed over time is hence ignored as well as direct increase in DDS. The effect of using DDS from the TIMER-travel dataset can be seen in Figure 69, right graph,

compared to constant DDS, left graph. It can be seen that the TD for HSM becomes twice as in the right graph. This model is the only model that is this affected by DDS gradients, and it is therefore decided to keep DDS constant for the evaluation of the GCAM model.

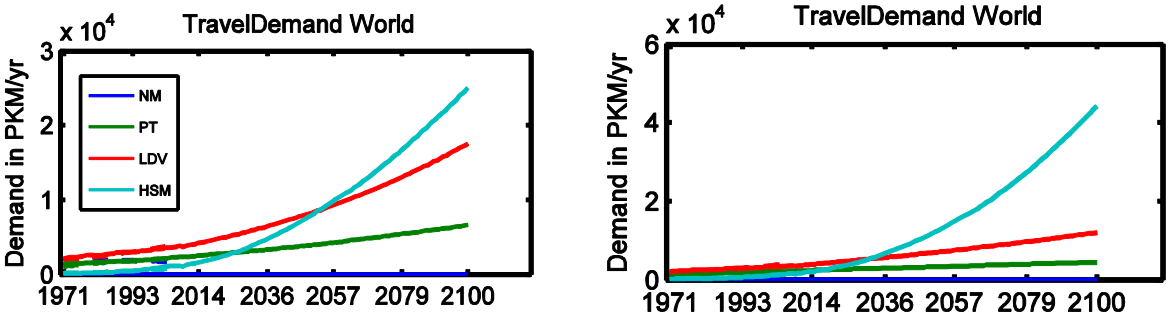


Figure 73: Left graph is a projection with constant DDS as in the original GCAM formulation. The right graph is without constant DDS as in the original TIMER-travel vehicles module. GCAM

3.3.3 RANGES

In determining the set of ranges for the Monte Carlo analysis, three independent variables are used. The solutions are strongly dependent on these variables and it is therefore decided to first analyze the least sensitive variables, see the sensitivity analysis in the previous section, and then use their values in order to determine the next variables. The cost distribution parameter μ, the income elasticity ϵ and the price elasticity σ.

GCAM uses a single elasticity for all modes, this makes the interpretation of their values more complicated than for SIMPLE or POLES, see section 0. σ is found to have a range from -0.6 to -1 independent of the mode. ϵ differs greatly between modes, increasing income I will increase TD of LDV and HSM and decrease use of PT and NM. Values for ϵ are hence found to be 0.2, -0.6, 1.0, and 2, for NM, PT, LDV and HSM, respectively. These elasticities need not represent the model correctly, however, as modal differences are facilitated by price competition.

It is therefore decided to use a value provided by the authors of the model. Which is the same as used in literature as a general elasticity (Schafer, Heywood, Jacoby, 2010), 1. The range is determined based on the size of the error found in literature, 0.3. This gives an initial range of 0.7-1.3. Lastly μ cannot be derived directly from literature, but has been provided in direct communications with the authors to be -2.

Depicted in Figure 74 is the dependency of TDfit on μ it can be seen that the TDfit is barely affected. This is due to the method of calibration, the relative values of SW are related to each other by $SW_{m1}^{\mu} \propto SW_{m2}^{\mu}$, for some mode m1 and some mode m2. The SW per mode can hence be scaled by some factor of μ and the error, TDfit, will remain identical. For the future projections the parameter does have a significant impact, see Figure 75. Due to this ambiguity the value provided by the authors +/- 10% is used.

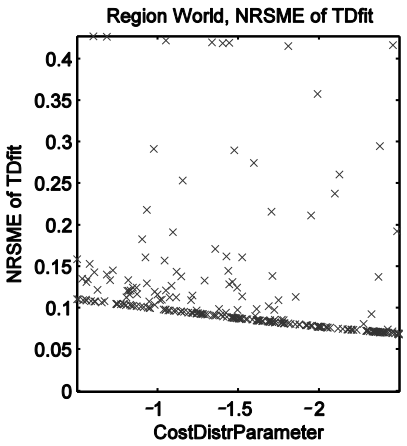


Figure 74: Sensitivity of the error, TDfit, on μ. GCAM

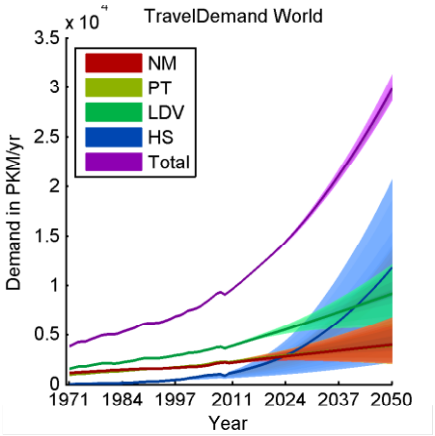


Figure 75: Impact on historic (<2005) and projected (>2005) TD when changing -μ over a range 0.2-4. GCAM

Using said value for μ, ϵ is analyzed. Depicted in Figure 76 it can be seen that the model is sensitive to changes in the ϵ and would have the smallest error at a value of 1.3, the upper limit of the literature value. It is decided to analyze the elasticity per country, depicted in Figure 77. It can be seen the ϵ for the regions with the best TDfits found earlier (i.e. not FSU, MEA, CPA), is consistent and averages around 1.1. This fits literature and a range of 1.1 ± 0.2 is assumed.

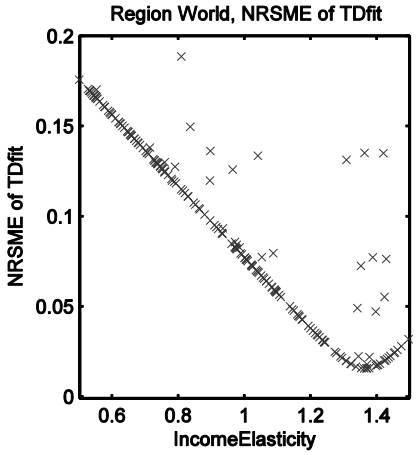


Figure 76: Sensitivity of the error, TDfit, on ϵ. GCAM

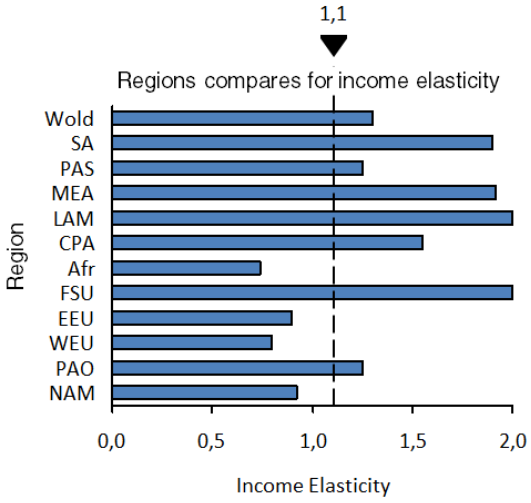


Figure 77: ϵ resulting from optimization per region ceteris paribus. The lack line is the median literature value. GCAM

Using the values for ϵ and μ, σ is plotted against the error, TDfit, for the world region in Figure 78. It can be seen that σ has smallest error for a value near -0.4, this is low considering literature. Therefore σ is plotted for each of the 12 regions, as depicted in Figure 78. It is clear from the deviation that σ cannot be derived accurately from historic facts. This can be explained by the strong dependency of the TD on income growth, price effects are much less significant. It is therefore decided to take the found literature value of -0.75 ± 0.25.

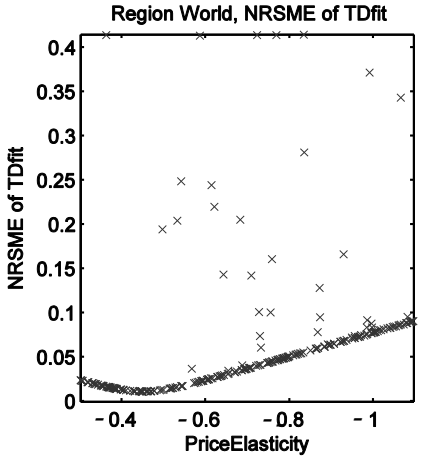


Figure 78: Sensitivity of the error, TDfit, on σ. GCAM

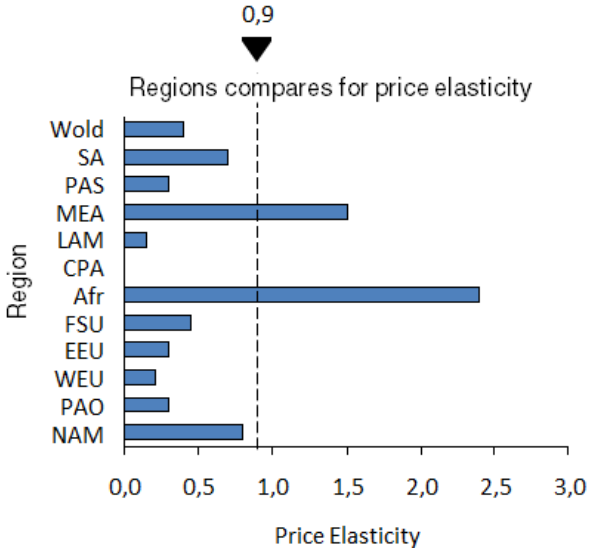


Figure 79: σ resulting from optimization per region ceteris paribus. The lack line is the median literature value. GCAM

A summary of the ranges found is given in Figure 80.

Independent variable	Symbol	Range
Cost distribution parameter	μ	-2 ± 0.2
Income Elasticity	ϵ	1 ± 0.3
Travel Price Elasticity	σ	-0.8 ± 0.2

Figure 80: Determined range per variable for the GCAM model

3.3.4 PROBABILISTIC PROJECTIONS

Probabilistic projections can be found in Figure 81. It can be seen that the TD for HSM increases rapidly has a high spread in Africa. The reason for this is the high rate of income growth, the spread is likely to be large. The TD in WEU doesn't increase as rapidly, TD for NM and PT remain constant.

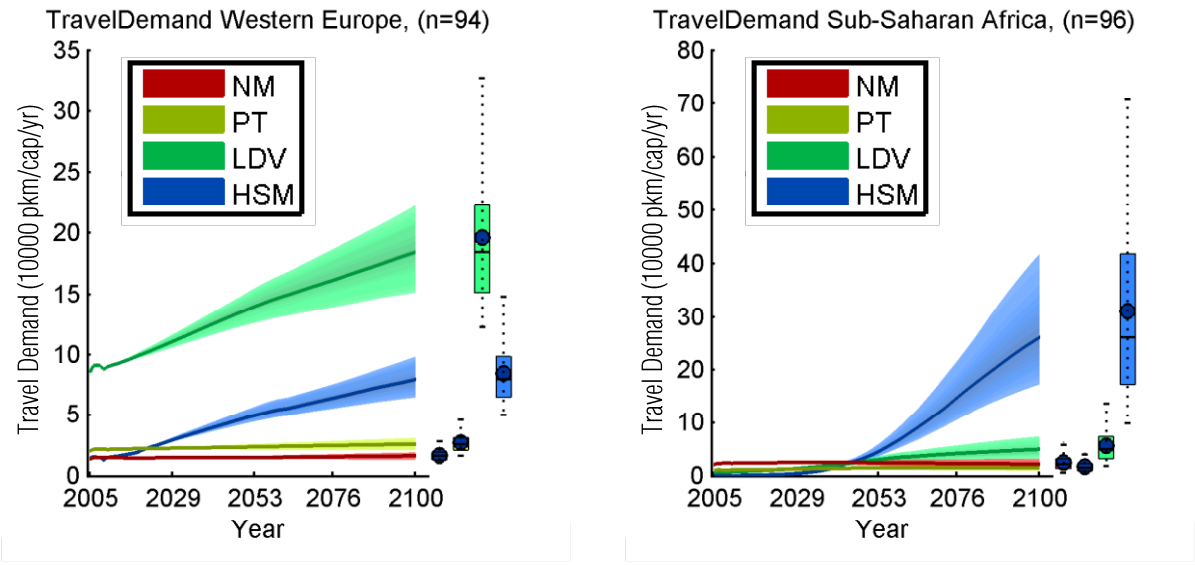


Figure 81: Select probabilistic projections. GCAM

Projections for the aggregated regions can be found in Figure 82. The indicated range shows a lower TD scenario is plausible. LDV development remains strongly present in the industrialized region. The development of TD for HSM in developing region has a large spread, the median scenario seems to forecast a large growth in the TD for HSM, with signs of saturation near 2100. For the reforming region the development is varied, and predicts a mix of all modes in 2100. TD for PT has a decreasing trend for all of the regions presented here.

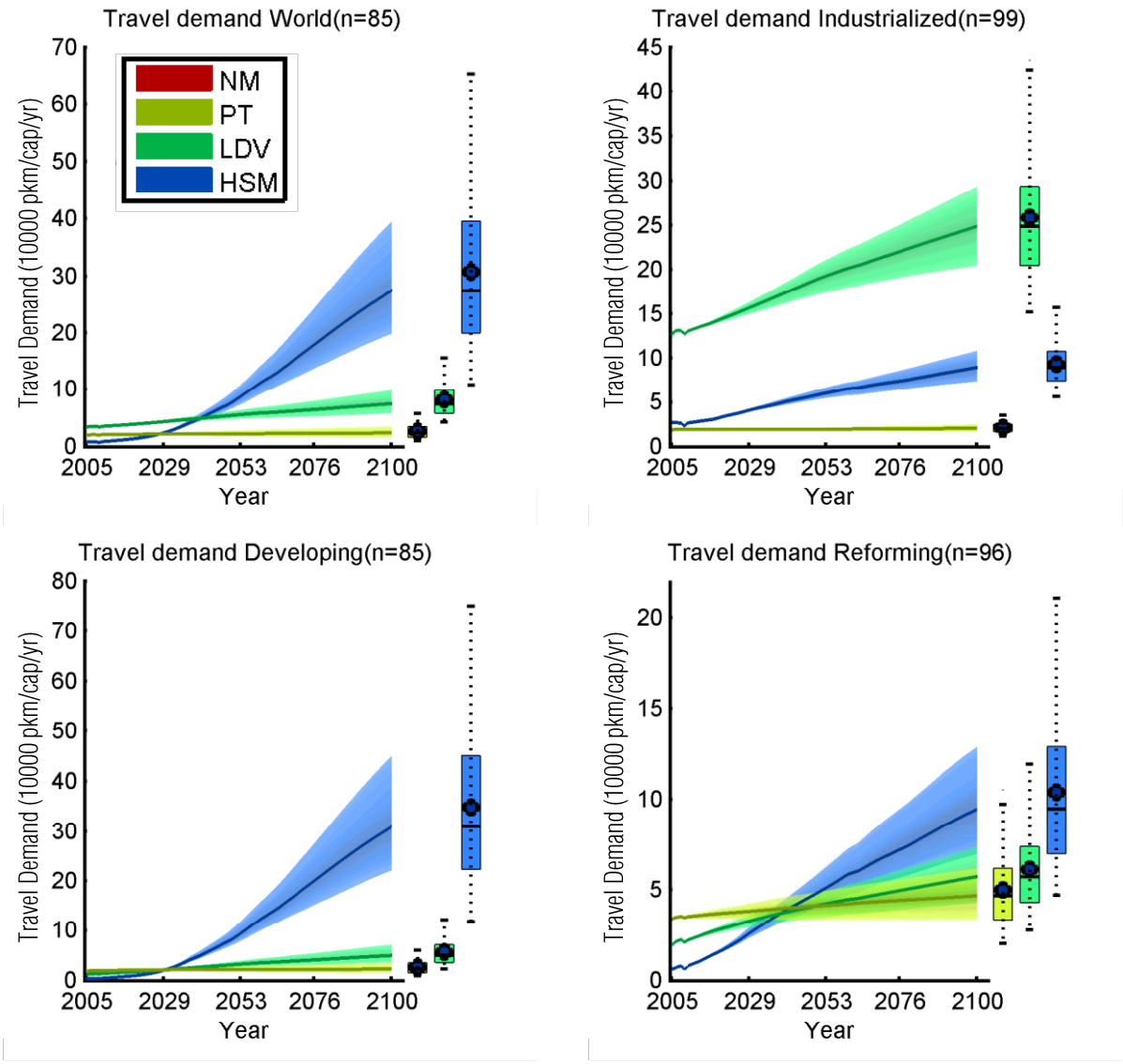


Figure 82: Probabilistic projections for the aggregated regions. GCAM

3.4.1 VALIDATION

The original TIMER-travel model has been modified slightly to include a sloped preference factor and a robust method for determining the time weight. Nonetheless the reconstructed model looks very similar to earlier publications, see Figure 83 (Bastien Girod et al., 2012).

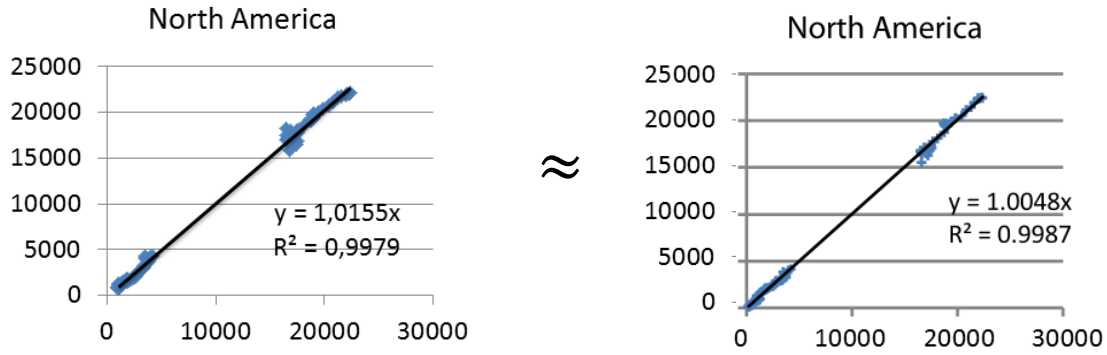


Figure 83: Comparison between reproduced (left) and published (right) values (Bastien Girod et al., 2012). TIMER-travel

An overview of the development of dependent variables for a model run is presented in Figure 84. It can be seen that the time weight TW stays near its initial value and that the normalized mode cost MCR for the modes converges in the future. The by eye fit of the mode split MS is as good as the fit by eye of the travel demand TD. Some minor deviations from the historic data can however be seen for LDV in 2005, the reason for this is probably volatility of the fuel prices.

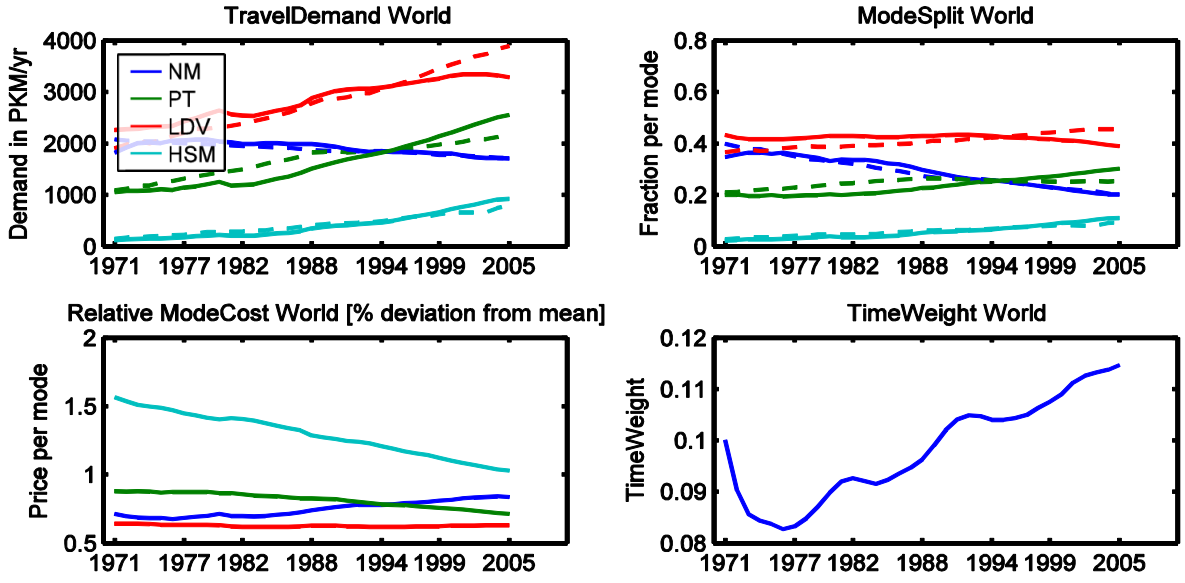


Figure 84: Historic travel demand based on calibration from 1971 to 1980 and projection from 1981 to 2005. The dotted lines are historic data. the TDfit is 2%. TIMER-travel

Presented in Figure 85 is the historic forecast for selected regions. The TDfit is well below 10% and therefore acceptable. The only deviation is for Latin America, the LDV mode grew faster in the historic dataset than forecasted by the model. This could either be data quality or poor interpretation of the dynamics of this developing country. There is reason to believe developing region can become motorized faster than predicted in this model, for example by buying motorcycles instead of cars (Chamon, M. Mauro, P. Okawa, 2008), see discussion.

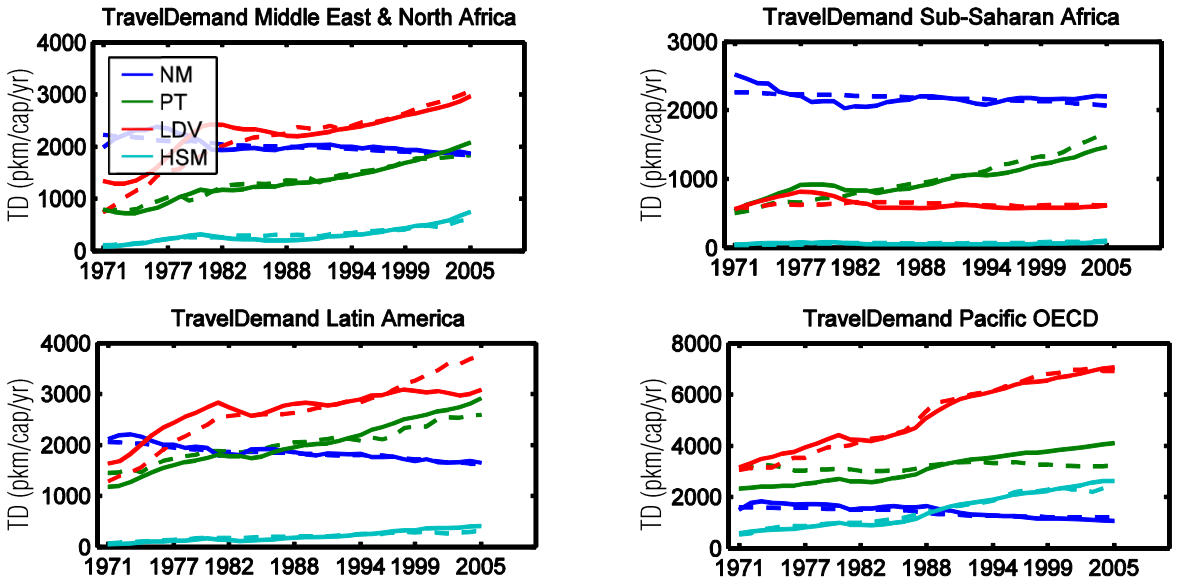


Figure 85: Historic travel demand based on calibration from 1971 to 1980 and projection from 1981 to 2005. The dotted lines are historic data. From left to right the errors, TDfit, are, from left to right, 5.8%, 5.9%, 3.4%, 3.9%. TIMER-travel

The error, TDfit, per region is next investigated in Figure 64. These errors are generated from the probabilistic projection runs using the ranges in the mentioned earlier. As clarified earlier a reasonable fit might be expected with a value lower than about 10% (Ruijven et al., 2009). Using this as an indicator, it becomes clear that all regions are valid except China (CPA), which has no run which fulfill this requirement and can therefore valid.

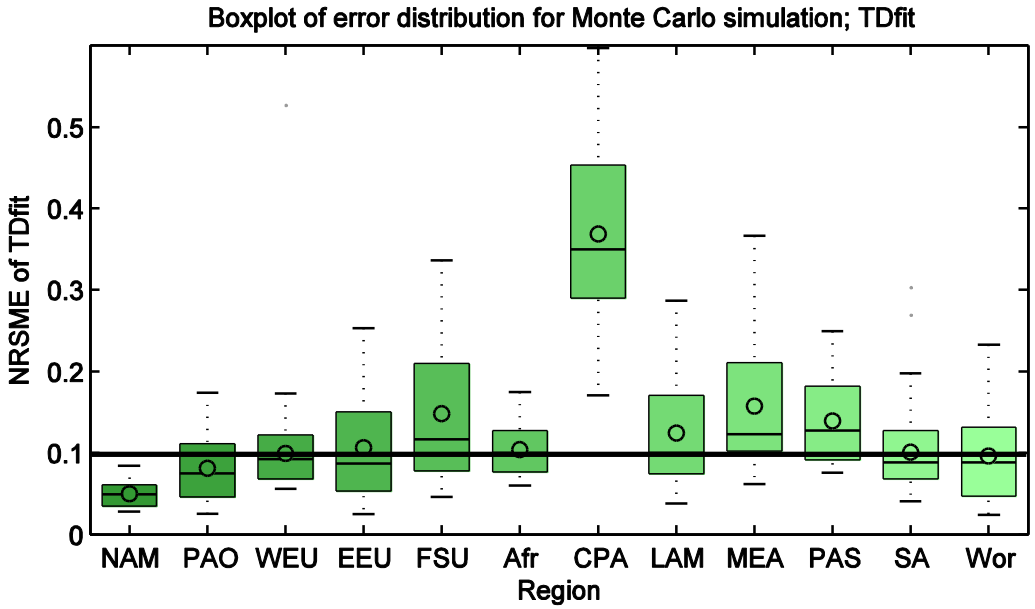


Figure 86: Boxplot of the TDfits over the unfiltered range used in the Monte Carlo analysis per region. The line indicates the threshold for the filtering. TIMER-travel

3.4.2 ANALYSIS

The sensitivity analysis for the Preference factor start PFstart is plotted in Figure 87. For HSM higher PF for HSM increases travel demand TD whilst for all other modes high PF decreases TD. This makes sense in light of the fact that more income spent on TD for HSM leads to higher TD than any other mode. Sensitivity for the remaining variables is depicted in Figure 88. It is remarkable that a higher travel time budget TTB leads to lower TD, the TMB thus has to be spent more will hence be spent on slower modes. Higher ψ leads to lower TD, because more money will be spent on comfortable travel, rather than TD. The model is not sensitive to the remaining variables.

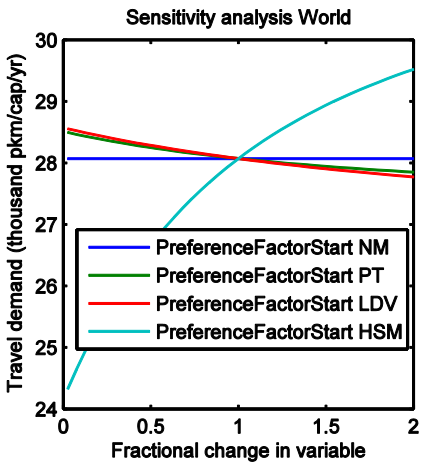


Figure 87: Sensitivity analysis, non-normalized values are 10000, 1, 4 and 2 of the PFstart for NM, PT, LDV and HSM, respecstively. TIMER-travel

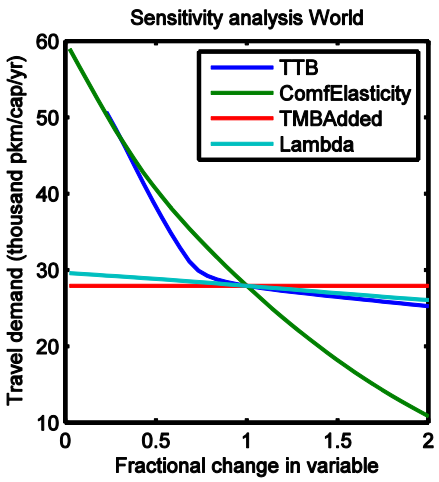


Figure 88:Sensitivity analysis, non-normalized values are 438,0.5,0.01 and 3.2 to the CF (TMBadded), ψ, TTB, and λ , respectively. TIMER-travel

The Total Travel Costs TTC over time is displayed in Figure 89. It can be seen that NAM has a maximum TTC 4% higher than any other region. CPA has a small TTC explained by the lacking development of HSM for this region and the low value compared to the TD of NM is explained by a high Travel Money Budget TMB correction factor (CF) needed for the initial value of the US. The TMB for this model is always equal to the TTC, because this is a boundary condition.

The Time Use TU is adhered to quite well, as can be seen in Figure 90. This makes sense in light of this being a loose constraint (Bastien Girod et al., 2012). CPA doesn't adhere to the TTB, due to model instability and therefore not considered valid as a region.

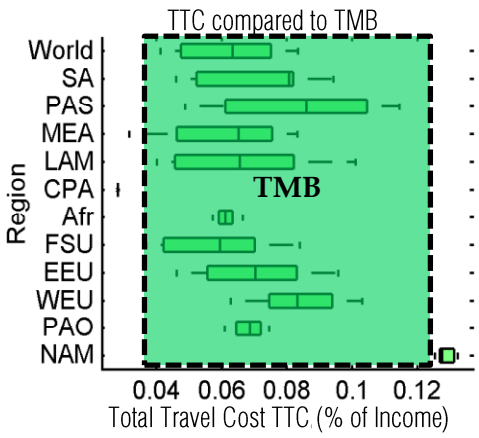


Figure 89: Boxplot of TTC for 130 years for the best Monte Carlo run of all regions. Dashed box represents the range of found values of TMB. TIMER-travel

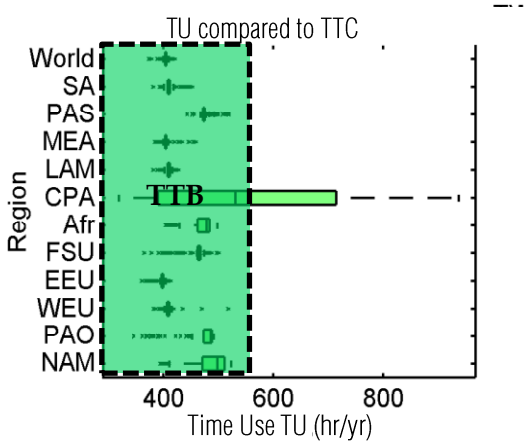


Figure 90: Boxplot of TU for 130 years, all modes, for the best Monte Carlo run of all regions. Dashed box represents the range of found values of TTB. TIMER-travel

The normalized mode cost MCR per region for LDV and HSM are plotted in Figure 91 and Figure 92, respectively. It can be seen that the MCR for LDV differs across regions, but not over time. Only for SA and PAS is LDV more expensive than average per pkm compared to the other modes. The HSM on the other hand develops into a relatively cheap mode for all regions, considering time develops the price from high to low for HSM.

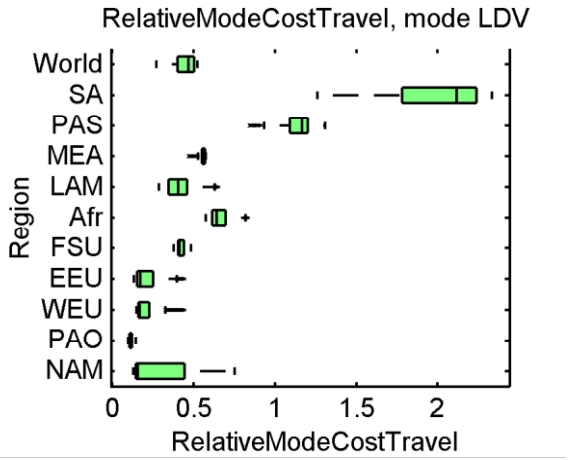


Figure 91: Box plot of relative mode costs, LDV, the range of values represents the change over time

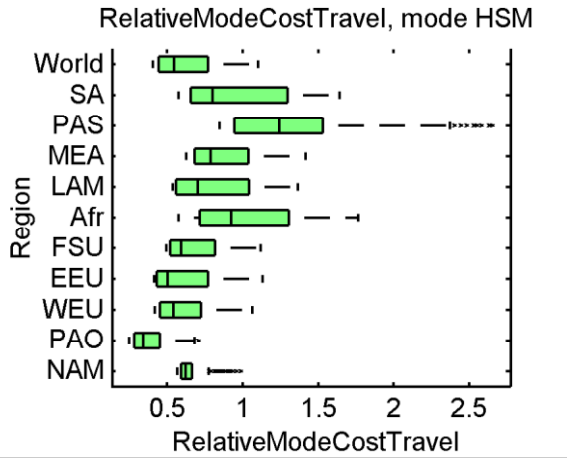


Figure 92: Box plot of relative mode costs, HSM, the range of values represents the change over time

The value of PF as defined in the original publication (Bastien Girod et al., 2012) is determined by calibration and fixed. Because preferences can change over time it is likely that the PF will change in the future. It might therefore be beneficial to include PF that is not fixed over time. Results are displayed in Figure 94 and Figure 93 for constant and sloped PF. The sloped Preference factor gives a superior (TDfit of 0.060 compared to 0.065) fit. Both are acceptably small, but a difference can be identify visually. The computational costs of this modification are minor, and it is therefore decided to use this in the final model runs.

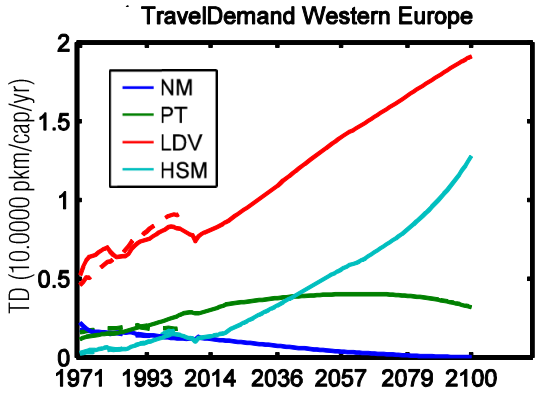


Figure 93: Projection with a fixed Preference factor, historic data is dashed. TIMER-travel

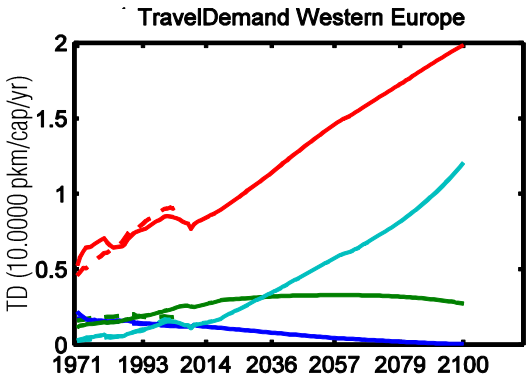


Figure 94: Projection with a sloped Preference factor, historic data is dashed. TIMER-travel

Values for the PFstart are displayed in Figure 95. They don't differ much from a recent publication, although they are defined inversely in that publication (Bastien Girod, Van Vuuren, et al., 2013). The PFstart values indicate clearly that for industrialized region LDV is preferred

and in developing region HSM. This finding is one of the reasons a convergence to industrialized preferences is added to the models. Values for PFslope are given in Figure 96, it indicates that the PF for PT and HSM don't increase much but NM and LDV do.

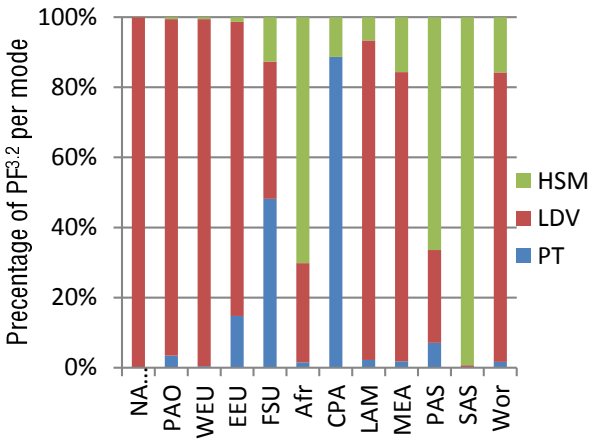


Figure 95: PFstart for all regions, average result of 250 Monte Carlo simulations. TIMER-travel

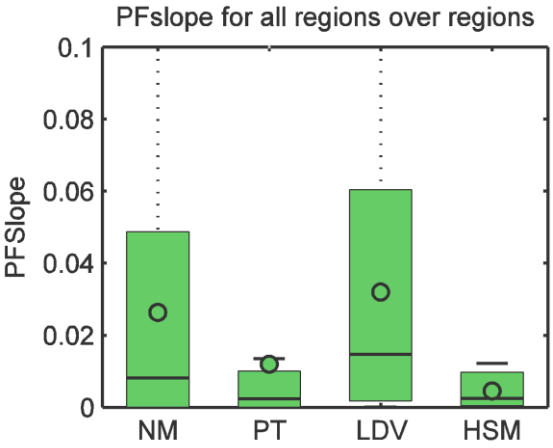
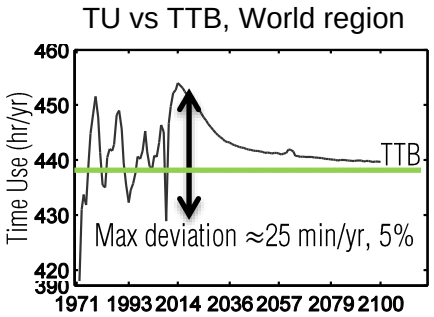


Figure 96: Range of PFslope for all regions. TIMER-travel

Unlike the other models in the present investigation TIMER-travel restricts the TTC to exactly TMB and the TU is loosely bound to TTB. The TW can alter the TU in order to ensure it converges to TTB, by adjusting relative prices,. In order to get TU to approximate TTB two methods have been published. Firstly the TW can be iterated over multiple times per year in order. Second, the entire model can be optimized using a nonlinear optimization algorithm to reduce the difference between TTB and TU.

The iterative solution is depicted in Figure 97 and Figure 98, with one iteration per year and five iterations per year, respectively. These solution with five times as many iterations is logically five times as computationally intensive. The optimization solution has also been created, it fits the TTB perfectly but is computationally hundreds of times as intensive as the iterative approach. The TTB has a broad range of values from 1.0 to 1.5hr/day, there is therefore no reason to believe a small deviation for 1.2 is unrealistic. It is therefore decided to use the iterative solution with as many iterations as years.



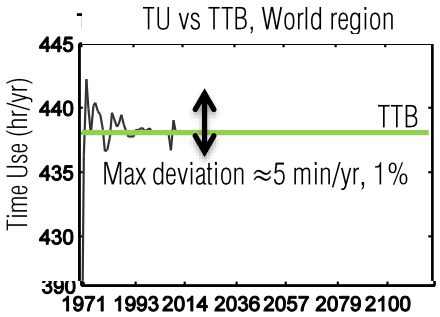


Figure 97: TU plotted over time, the black line represents the median TTB value. Model is run without any additional iterations. TIMER-travel

Figure 98: TU plotted over time, the black line represents the median TTB value. Model is run 5x as many iterations. TIMER-travel

Unstable regimes exist that relate to the TW. They are created either when two modes compete see or when time weight approaches zero and TW has no influence on the model. This can be seen in Figure 99. The issues with competing modes are rare and unavoidable and are left to be rejected by the normal post-processing methods. The issue of low time weight is solved by not allowing the time weight to reach a value lower than 0.1. A minimum of 0.1 is taken because empirical testing shows this to create stable regimes for all regions. Because TW is only used in order to satisfy TU ≈ TTB it is easily tested if it is still effective, at a limit of 0.1 it is.

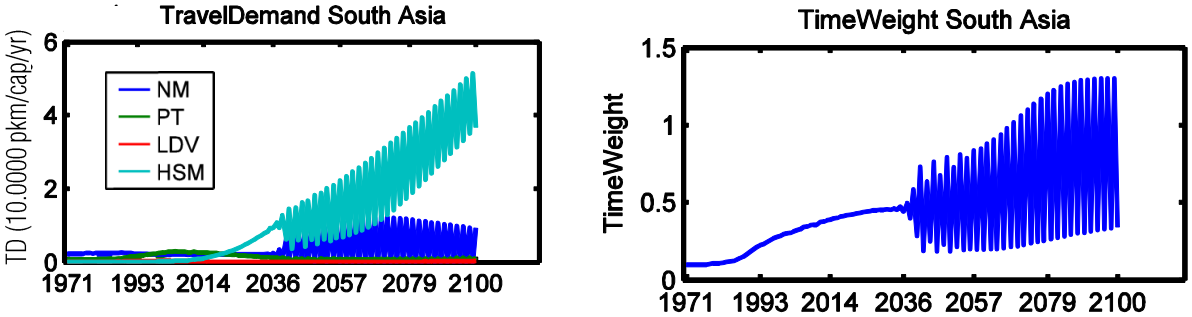


Figure 99: South Asia in the unstable regime for t>2035, TD and TW for the same run, left and right, respectively. TIMER-travel

3.4.3 RANGES

In determining the set of ranges for the Monte Carlo analysis, five independent variables are used. The TMB correction factor CF, Lambda λ, the Travel Time Budget TTB, Comfort Elasticity ψ and Inertia α.

The analyses are presented in Figure 100. Each model run varies only the variable depicted and optimizes over the Preference factors as usual, over the whole historic set, 1971-2005. It is easy to see that adjusting the travel money budget quickly yields erroneous fits for the historic fit. These ranges are kept below the 10% error margin, hence 0.5 ± 0.1 for ψ.

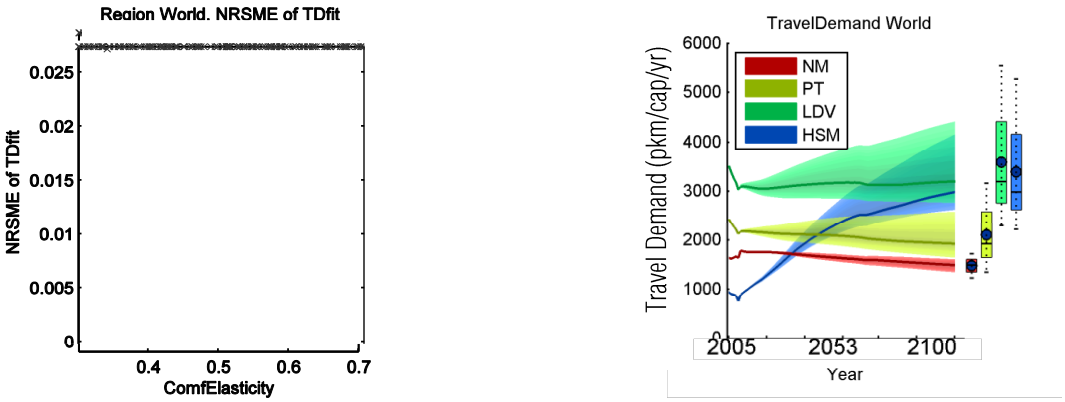


Figure 100: Sensitivity of the error, TDfit, on ψ and a probabilistic projection varying ψ from 0.3 to 0.7, ceteris paribus. TIMER-travel

The TMB added factor TMBA is taken to be 0% ± 1%, similar to literature values on the uncertainty of TMB (Schafer & Victor, 2000). α variable appears very insensitive and will therefore be ignored altogether, it's value 1.The TDfit is insensitive to the TTB, but the model appears to undergo a behavior change for values >500, the behavior above that value might no longer be behaviorally correct. The value is therefore allowed to change in a broad range from 438 ± 50 hr/yr.

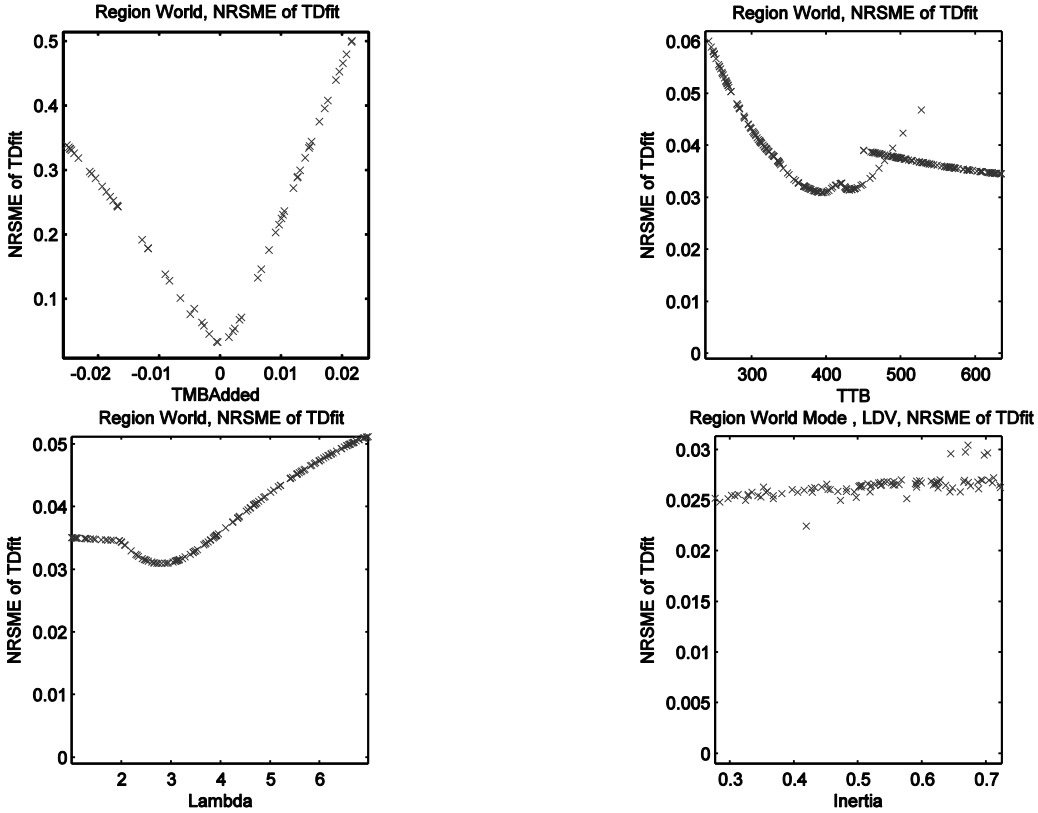


Figure 101: Sensitivity of the error, TDfit, on, from left to right TMB, TTB, α, λ. TIMER-travel

The Final results for the ranges used in the probabilistic projections can be found in Table 15.

Variable	Symbol	Range
TMB Added	TMBA	0 ± 1%
Lambda	λ	3.2 ± 0.5
Travel Timer Budget	TTB	438 ± 50 hr/yr
Comfort Elasticity	ψ	0.5 ± 0.1
Inertia	α	1 (ignored)

Table 15: Ranges used for the probabilistic projections for TIMER-travel

3.4.4 PROBABILISTIC PROJECTIONS

Probabilistic projections for sample regions can be found in Figure 102. For WEU the large initial TD for LDV will remain throughout the century and only begin to overlap with HSM near 2100. Africa is predicted to skip LDV as a dominant form of transport and move from dominant PT to dominant HSM. In Africa the TD for HSM is larger than any other mode by 2100, although the simulation has a fairly large spread.

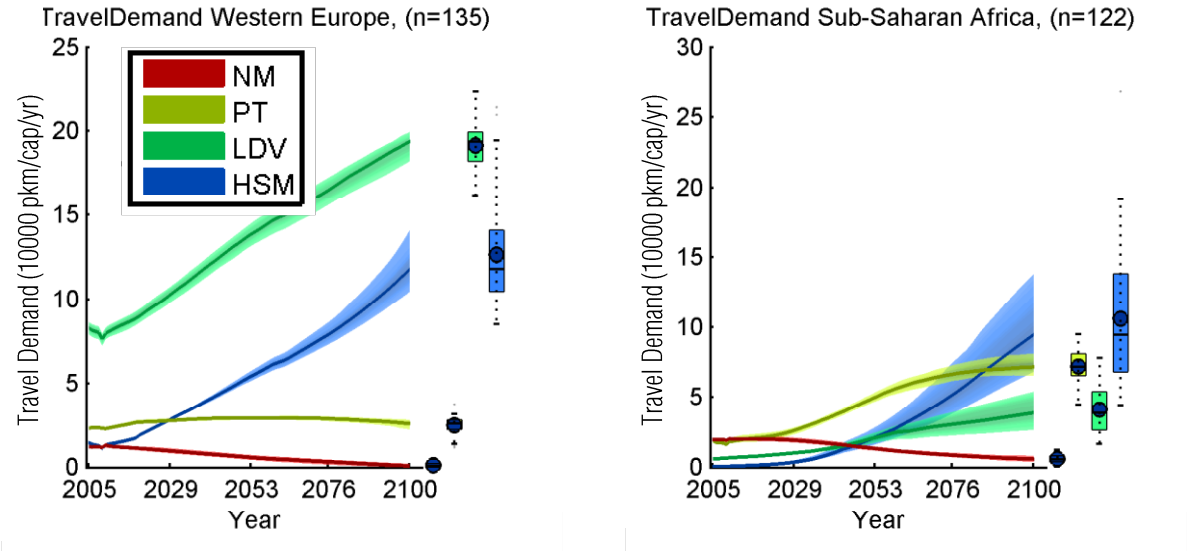


Figure 102: Select probabilistic projections. TIMER-travel

The aggregated probabilistic projections based on the TIMER-travel model can be seen in Figure 103. The worldwide trend is clear, away from non-motorized NM (left out here for comparative reasons) and PT toward LDV and very strongly towards HSM. By 2035, in fact, HSM is predicted to be the largest mode, for the median solution and by 2100 it can be seen that all scenarios point to HSM being between two or three times as large as LDV, with no saturation for either modes by 2100. This is mostly due to the developing region which had a very large share of HSM. For reforming and developing region there is a PT mode that is almost as large as LDV, and only starts to decline later in the century.

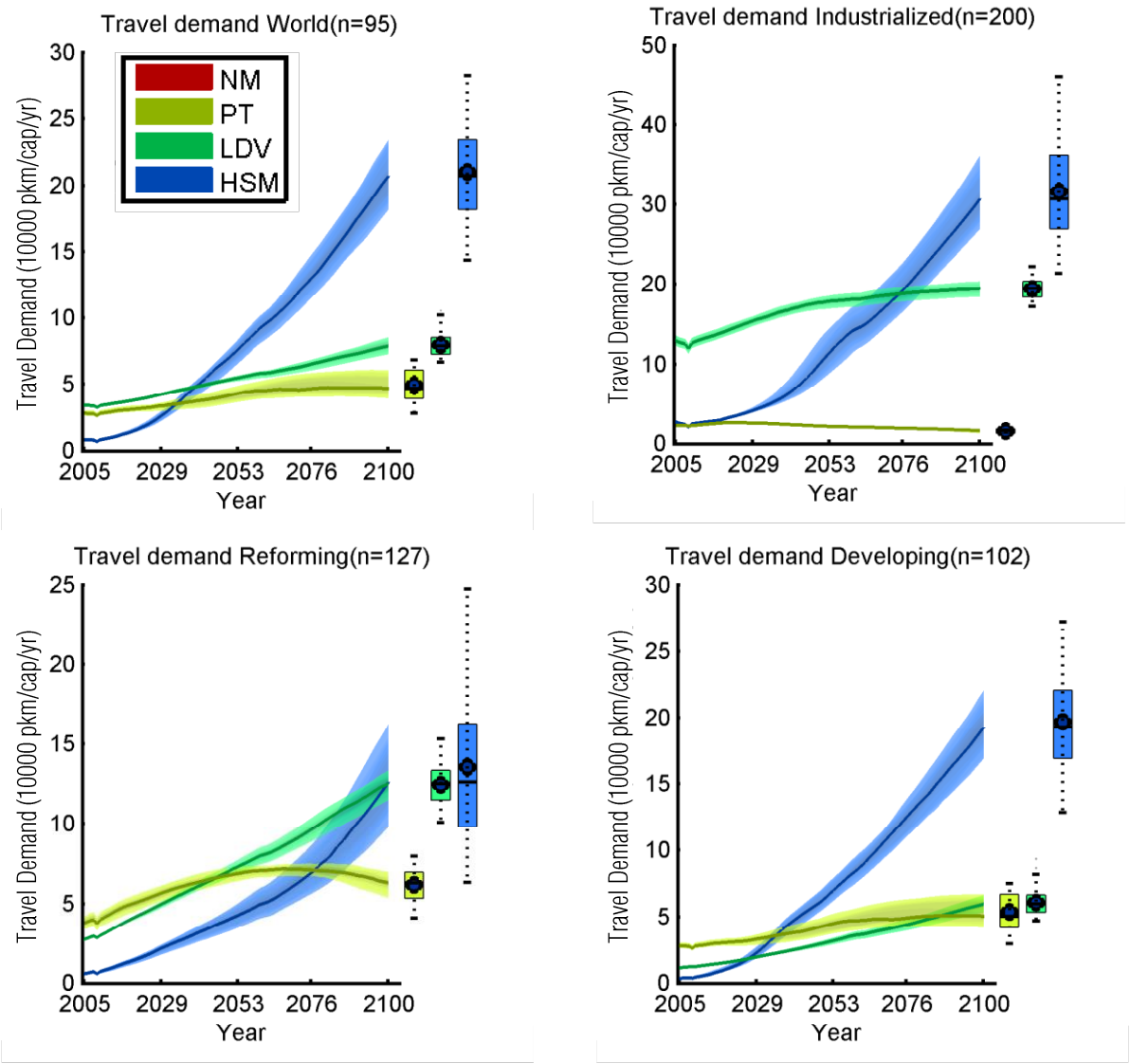


Figure 103: Probabilistic projections for the aggregated regions. TIMER-travel

3.5 COMPARISON OF THE MODELS

This section contains the comparisons of results. First the analyses for the models will be compared, these exclude the probabilistic projections. Next the probabilistic projections will be presented as they were, without added convergence of the world to the industrialized region. Lastly the same results will be presented with convergence of preferences to the industrialized region.

3.5.1 INDIVIDUAL ANALYSES COMPARED

An overview of the models is presented in Table 16. It can be seen that the models differ greatly in their ability to both recreate historic datasets and fulfill boundaries of TU ≈ TTB and TTC ≈ TMB. GCAM and TIMER-travel perform well, with exception of China (CPA) which has a large error, TDfit. For POLES, however the reforming and developing region have to be rejected on basis of the historic fit. For SIMPLE the same rejection applies to the developing region. From top to bottom the rows of the table represent:

- The number of independent variables can be read, and which of these variables are calibrated.
- Variables which are regionally dependent are depicted for each model. Due to the regional definition of the independent variables, which make the number of variable very large, for POLES and SIMPLE ambiguity exists. All these variables describe the TD, and this means the models are over specified, which leads to ambiguity in input. For example, the Time trend functions identically to the ϵ in the POLES model, for a positive gradient on income over time, a distinction cannot be made during the optimization procedure.
- Two of the models, TIMER-travel and POLES, additionally have unstable regimes that have to be filtered out. This means automation is difficult for some regions and the model has to be adjusted until the instability is overcome.
- The adjustments presented are overcome by the adjustments mentioned. These are feasible adjustments, but were not described originally by the authors.
- The fit of the time use TU with the empirical finding of the travel time budget TTB
- The fit of the total travel cost TTC with the empirical finding of the travel money budget TMB
- The historic fit in terms of the TDfit, the general threshold is 10%, but depend on the model and region

Characteristic	TIMER-travel	GCAM	POLES	SIMPLE
Independent variables of which (calibrated)	8 (5)	8 (5)	13 (3)	8 (0)
Independent variables worldwide?	Yes	Yes	No, 11 regions	No, 11 regions
Unstable regimes	Yes, rapid jittering of TD and TW	None	Yes, TT causes TD > TDSAT or ER > ERSAT	None
Adjustments	PFslope & restricted TW	Constant DDS, removal of ψ from TP data	Restricted range of TT	N/A
TU ≈ TTB	Loose boundary	Acceptable*	For ind only	TU ≈ $2x$ TTB
TTC ≈ TMB	Boundary	Yes *	For ind only	Yes for ind & ref
Historic fit, NRSME	<10% world*	<10% world*	<10% industrialized region	<10% industrialized & reforming
Accepted regions	World*	World*	Industrialized	Industrialized & reforming

Table 16: Summary of the individual comparisons of the models, * refers to the exclusion of CPA

In summary of the individual error analysis some regions have been rejected for several models. A summary of all the variables presented in the individual analysis section is presented in Table 17. The rejected regions vary per model. CPA is rejected for all region.

World											
Developing						Reform.		Industrialized			
SAS	PAS	MEA	LAM	CPA	Afr	FSU	EEU	WEU	PAO	NAM	
											TIMER-Travel
											GCAM
											POLES
											SIMPLE

Table 17: Accepted, green, and rejected regions for each model

3.5.2 COMPARISON OF PROJECTIONS; ORIGINAL MODELS

In comparing the probabilistic projections an overview of the different aggregated regions and their probabilistic projections are presented here for the aggregated regions. The TD has been summated across all modes. Probabilistic projections for Energy and CO_2 are omitted and can, with the recommendation not to take these at face value, be found in the appendix.

Figure 104 presents the aggregated probabilistic projections, for the aggregation to global regions the TD per capita of CPA is assumed to be equal that of the developing region. First looking at the worldwide TD, there is a remarkable level of agreement between the models on the median scenario and that there will be a 6 to 12 fold increase in worldwide TD (50% of simulations for GCAM). Compared to only a doubling in the industrialized and reforming region, most of the growth will come from the developing region, which are predicted to increase 8 to 16 fold in their total TD. If one looks at the distribution of the TD it is easy to see that there is a shift from TD distributed equally across developing and industrialize. The smallest values in the distribution for 2100 for the developing region is 0.25 Ppkm/yr compared to the largest in the industrialized region, 0.05 Ppkm/yr yields a minimal 5-fold difference between the two. In other words, the developing region will have five times the TD of the industrialized region.

Regarding the differences between the models there is no probabilistic projection significantly different in 2100 at the p=5% level using a standard T-test for the valid regions. Worldwide there is no difference in median prediction, only in spread. For industrialized region spread is similar and the prediction for 2100 is similar for three models, but different for SIMPLE. It is obvious, however, that the SIMPLE and POLES model differ from the GCAM and TIMER-travel model in that they do not have signs of saturation of the TD by 2100, in POLES this is in a large part due to high saturation in developing region. The GCAM and TIMER-travel model together provide similar predictions regarding the TD in 2100 and its' development worldwide. However, their predictions for the reforming and industrialized region the prediction differ, most notably after 2050 for the industrialized regions.

The spread of the projections for the developing region tends to be larger for GCAM compared to the other models. This is largely due to the rapid growth of the GDP and the relationship between GDP growth and TD growth. The ϵ is a range and therefore affects strong growths by producing a relatively greater spread than a model that does not directly depend on elasticities, such as TIMER-travel.

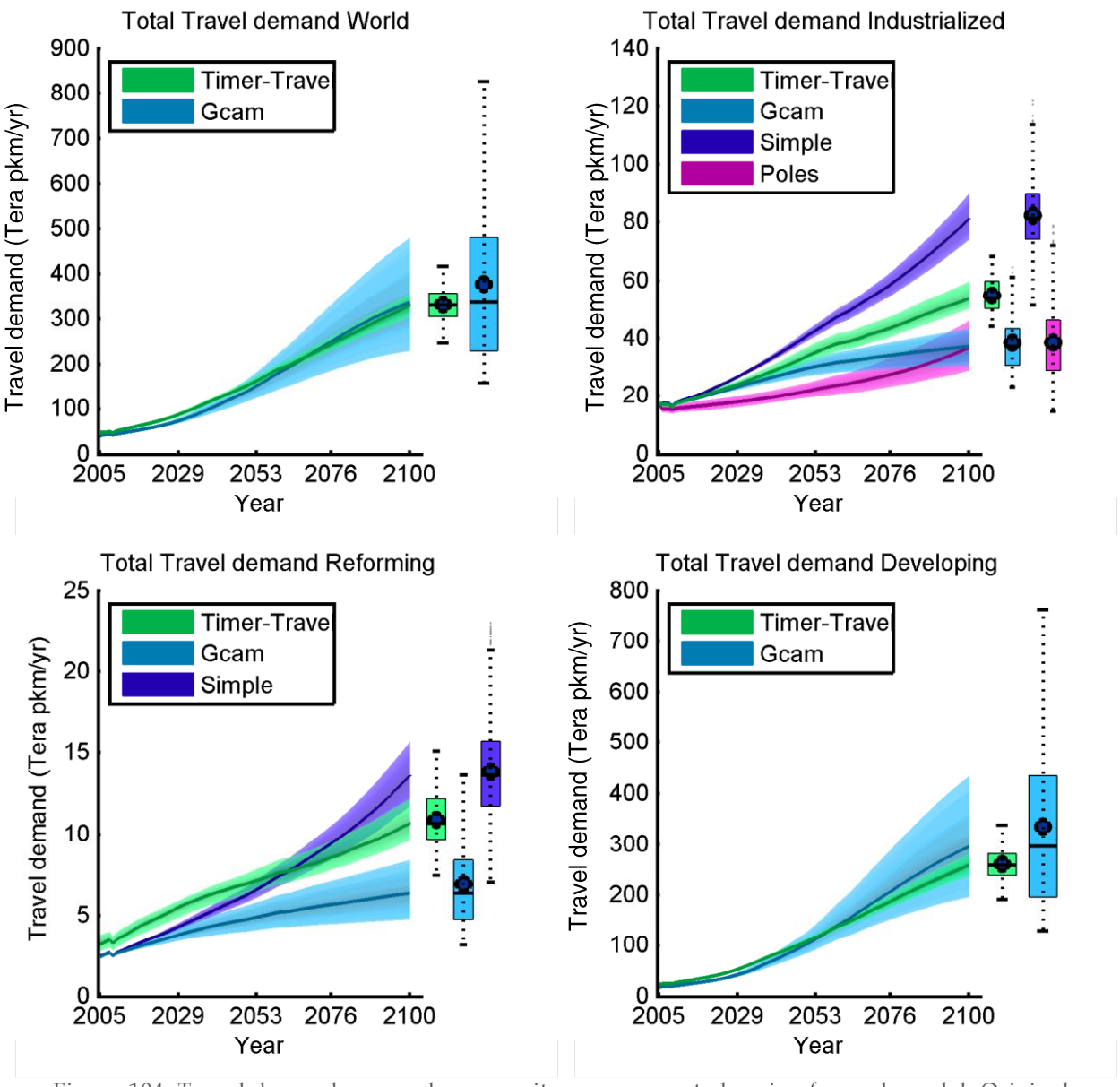


Figure 104: Travel demand summed over capita per aggregated region for each model. Original models

Figure 105 presents the time use TU for each of the models investigated. It can be seen that the TU for PT is much higher than one might infer from the TD data, this is because of the relatively low door to door speed DDS. The amount of time spent in HSM is only a fraction of that spent in other modes. It can also be seen that the POLES model in particular exceeds the TTB of 1.5 hr/day in reforming- and developing region. SIMPLE exceed the TTB for the industrialized region, however this might be due to the fact that the lowest car speed (60 km/hr) is used to create these graphs and the median run is taken, which might not represent the best model run.

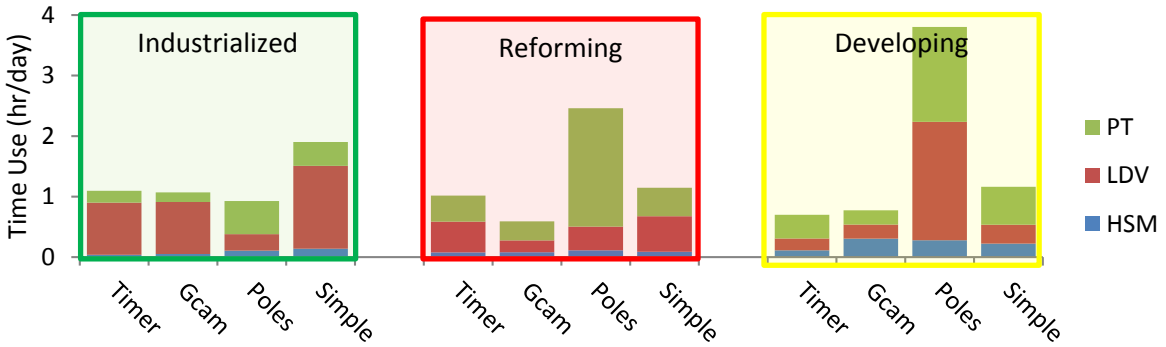


Figure 105: Time use per model, per mode, per region[19]. The median run in the probabilistic projections is used. Original models

World maps for the probabilistic projections for each of the models are displayed in Figure 106. The overview hence presented has largely been discussed in the previous section, with a few changes. Firstly it can be seen that for South Africa the TD increases more rapidly than any other region, the region has a large population (1.3bn), therefore this growth strongly affects worldwide TD. Secondly Africa also has a rapidly growing TD and has a spread is large beyond 2050. Lastly it can be seen that there is not a single model that gives the smallest or largest predictions for each region, rather it differs per regions.

[19] The GCAM model does not have speed development, HSM has a DDS of 270 instead of 490 in 2100

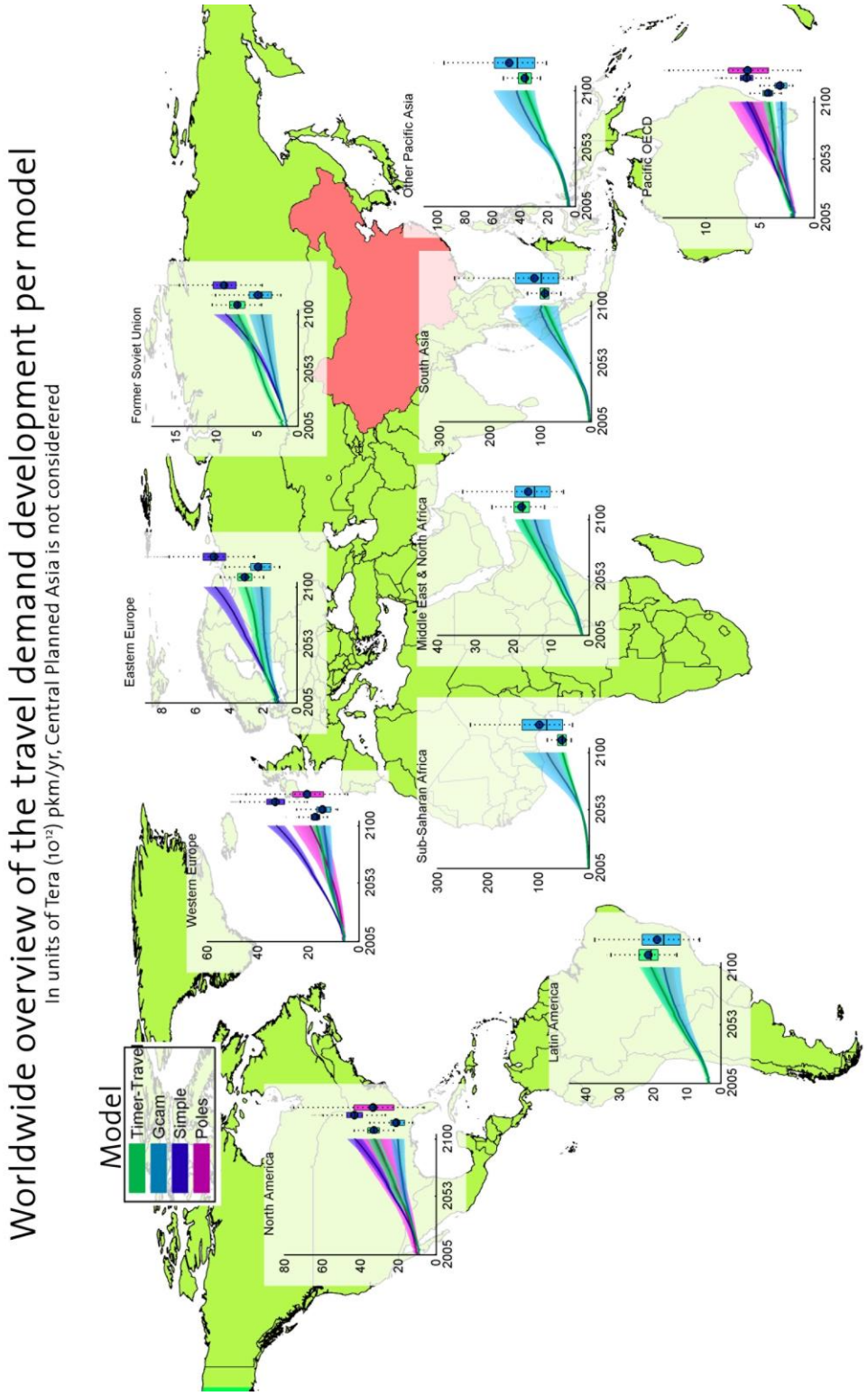


Figure 106: Probabilistic projections for Summated TD per region in units of Tera pkm/yr Original models

3.5.3 COMPARISON OF PROJECTIONS; WITH CONVERGENCE

Presented here are the results for the models with convergence of preferences to industrialized region added. The analysis is hence re-evaluated for GCAM and TIMER-travel, as described in section 2.3. This convergence mostly affects the developing and world regions. Probabilistic projections for TIMER-travel and GCAM are presented in Figure 108 and Figure 109. It can be seen that the TD for HSM is no longer largest. Instead TD for HSM and LDV have similar magnitudes. Comparing the two models to each other with a t-test does not indicate they are (at p=0.05 levels) significantly different. GCAM is, as before, the model with the larger projected TD and the larger spread.

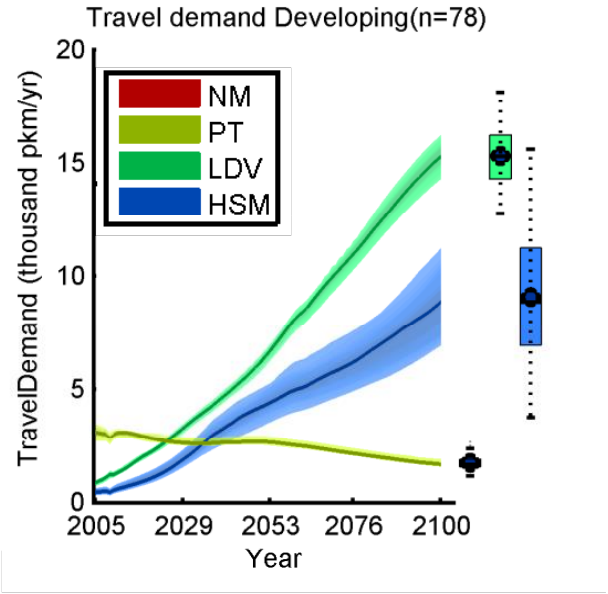


Figure 107: Probabilistic projection of travel demand per mode. TIMER-travel with convergence of consumer preferences

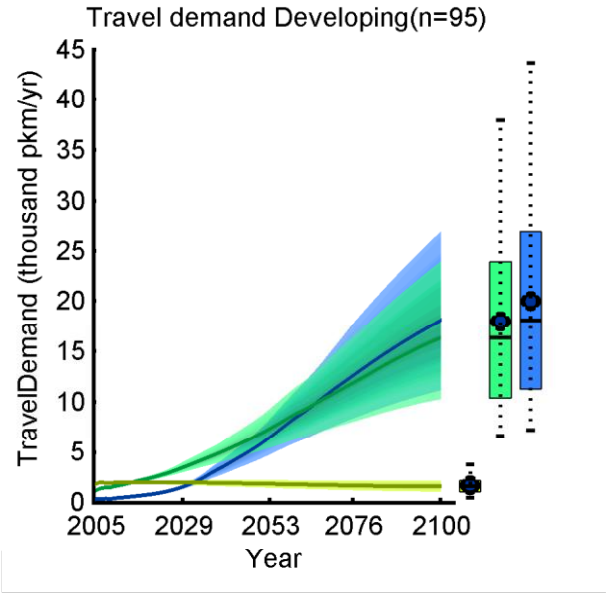


Figure 108: Probabilistic projection of travel demand per mode. GCAM with convergence of consumer preferences

Figure 109 and Figure 110 present the TD projections for the world region. It can be seen that the same conclusions can be drawn as from the developing region, with the notable exception that the HSM use seems lower for TIMER-travel, but not significantly

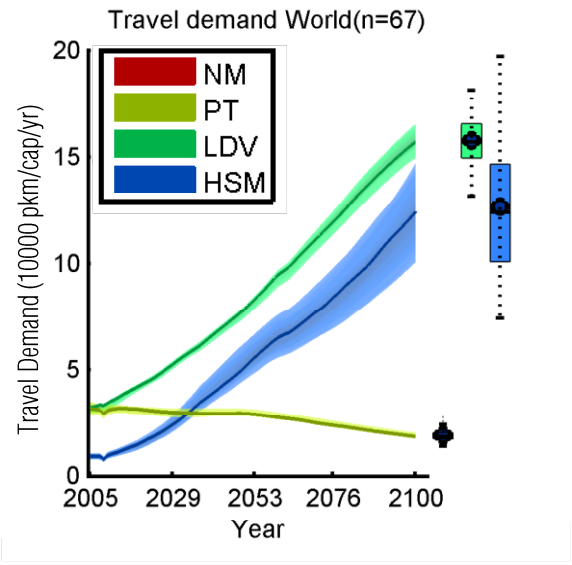


Figure 109: Probabilistic projection of travel demand per mode. TIMER-travel with convergence of consumer preferences

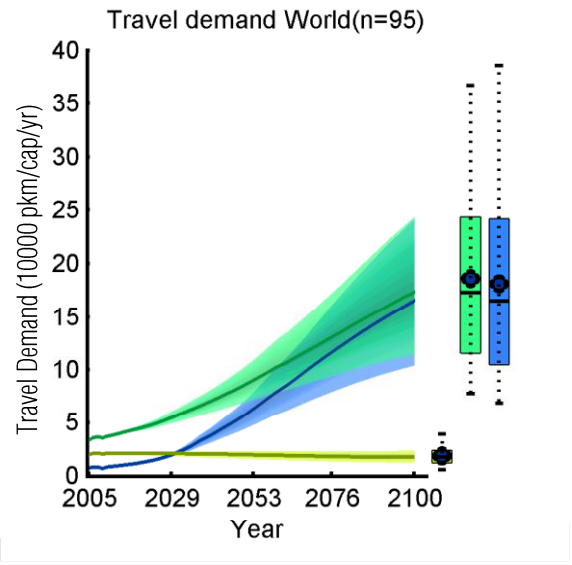


Figure 110 TD Probabilistic projection of travel demand per mode. GCAM with convergence of consumer preferences

Figure 111 and Figure 112 present the TD for the two models. It can be seen that the TD with is only slightly lower than for the original models. This is likely due to the weak coupling between mode split, travel prices and TD in both models. GCAM only feeds back through the pricing and TIMER-travel by the purchase of relatively slower modes of transportation.

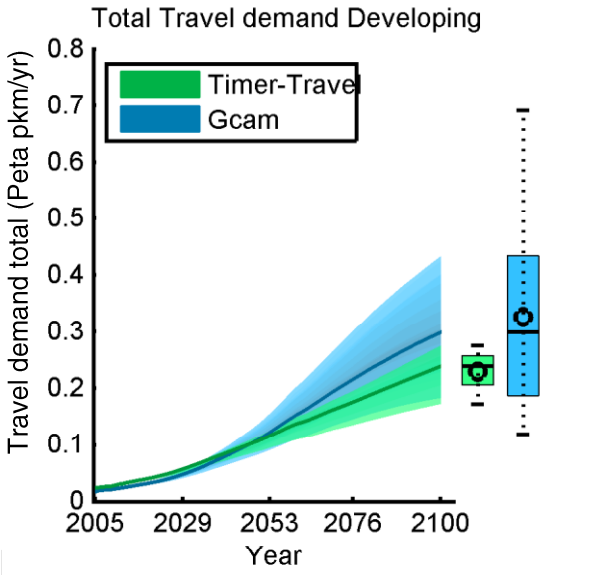


Figure 111: Travel demand summed over capita per aggregated region for each model. Models with converging consumer preferences

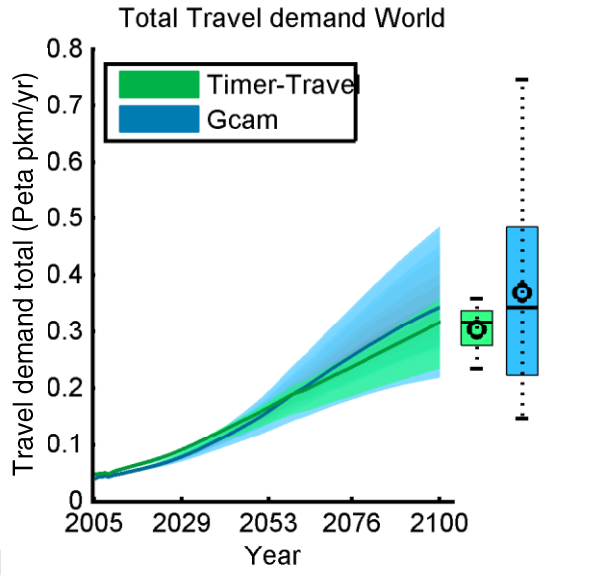


Figure 112: Travel demand summed over capita per aggregated region for each model. Models with converging consumer preferences

Figure 113: Time use per model without saturation modification. The median run in the probabilistic projections is used. Models with converging consumer preferencesFigure 106 presents the actual time use for each of the models. Only TIMER and GCAM changed from the previously presented graph due to saturations. It can be seen that in the Developing region the amount of time spent on airplanes is significantly reduced, and now much closer to the industrialized region's values.

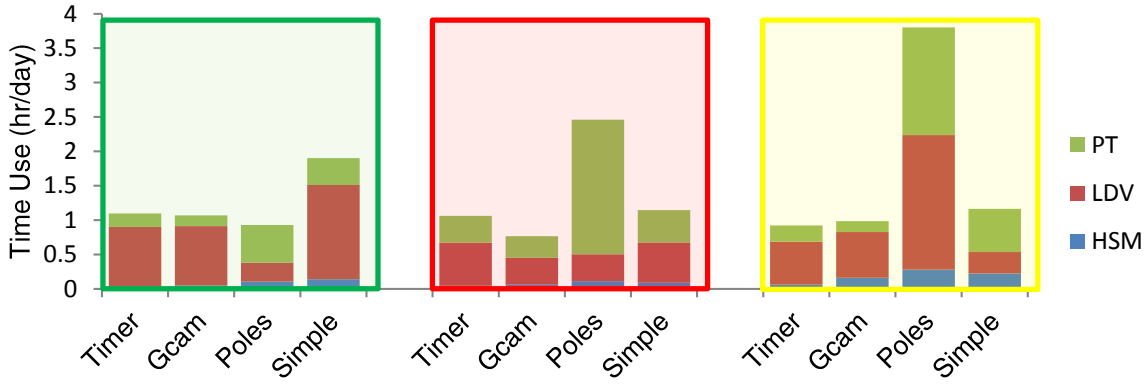


Figure 113: Time use per model without saturation modification[20]. The median run in the probabilistic projections is used. Models with converging consumer preferences

Figure 114 presented the overview of all regions, similar conclusions can be drawn as the runs without saturation. The POLES model has been included as a reference, because it natively contains saturation definitions similar to the convergence discussed, the non-industrialized regions for this model have to be interpreted with care.

[20] The GCAM model does not have speed development, HSM travels at 270 instead of 490 in 2100

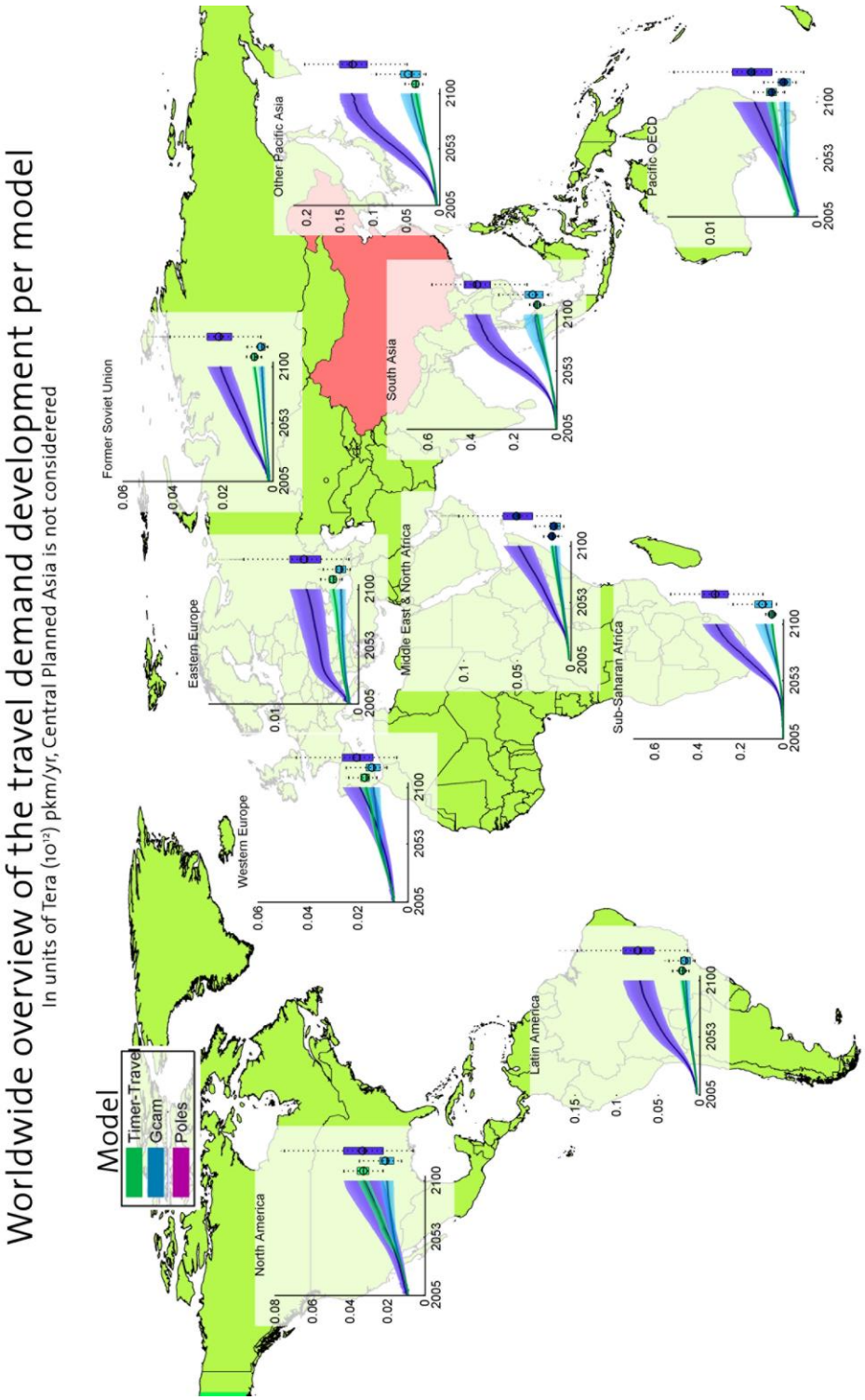


Figure 114: All regions' probabilistic projection of the summated TD compared. Converging preference modified models

4 DISCUSSION AND IMPLICATIONS

The discussion and implications are split up into four sections. The first section will compare the results for the given models to the literature sources, in order to prove or disprove their validity. The second section will outline implications of the present investigation according to the research questions posed, first for modelers, secondly with regards to the historic fits per model and lastly for the future projections made. The third will give shape to the limitations of the present investigation and implicate possible improvements. The fourth section will outline detailed future directions that logically follow from this investigation.

4.1 LITERATURE COMPARISON

The literature comparison compares the findings of other literature with the findings in the present investigation. It will first investigate the modeling methods used, secondly discuss the results aggregated across all modes, lastly it will discuss each mode separately.

4.1.1 MODELING METHODS

Regarding modeling methods, adherence to empirical findings, such as of the Time Use TU to Travel Time Budget TTB (Chen & Mokhtarian, 2008) and Total Tavel Cost TTC to Travel Money Budget TMB (Stopher & Zhang, 2011) have been investigated. It is found that for models that don't adhere to these empirical findings large errors in the historic fit, TDfit, exist. In this light it is remarkable that only the TIMER-travel uses these as boundaries or constraints in the model formulation. GCAM seems to follow the empirical facts regardless of explicitly formulating them. The POLES and SIMPLE models both fail to adhere to the empirical facts for the majority of their regions. Restricting the solutions of the models by imposing boundaries on TU and TTC might improve the results significantly, as has already been done in TIMER-travel. In determining the optimal solution one could be take the empirical facts as additional criteria for the optimization procedure (Sofer, 2009).

The elasticities in the models analyzed have been mode dependent and constant or constant with a saturation level for the SIMPLE and POLES models, respectively. These model did not forecast historic data correctly. They are simplistic because in the long run elasticities they have at least an income- , time-, and cross-modal dependency (Litman, 2012b), and there is also asymmetric behavior when slower modes are replaced by faster ones (Chamon, M. Mauro, P. Okawa, 2008). Elasticity based models are not inherently flawed as shown by the GCAM model, which also has elasticities but has its mode competition defined separately. Further investigation and careful definition of the dependencies of these elasticities might yield better results.

Modeling in general has multiple sources of error, the model, the data and the modeler. Ignoring errors made by the latter, the source of error can be somewhat identified due to the multi-model approach in the present investigation. As mentioned there is reason to believe that at least CPA has a poor dataset, as it doesn't fit well with any of the models. Additionally two of the four models did not forecast on historic data without large errors. This is then likely more due to the models themselves than to the data. Better sources for the regional values of variables could have

significantly improved the fits for these regions and models. Finally basing models on the industrialized region, or even the USA in the case of GCAM, as the models in this investigation do, leads to different dynamics in growing regions. Data from local studies in the region could possibly be combined to improve the sets (Zhang et al., 2007). This could, for example, include more detailed modes that are more appropriate to developing regions, such as biking or motor cycles.

4.1.2 AGGREGATED PROJECTIONS

Comparing the global TD projections to values in literature leads to similar results. Not many sources project up to 2100, so 2050 is taken for six models giving a range of 80-160 Tera pkm worldwide (Bastien Girod et al., 2013; Schafer & Victor, 2000), compared to 90-140 Tera pkm for the models in this investigation. Looking at forecasts for 2100 one finds 150 Tera pkm/yr for GCAM (Kyle & Kim, 2011) and 270 Tera pkm/yr for TIMER-travel (Bastien Girod et al., 2012), at the lower 25% of both models, which have a median prediction of ~400 Tera pkm/yr. Which is a little higher but explained by the skewed sensitivity analysis of the income elasticity. The prediction for 2100 is dominated by the developing region, the most uncertain region in terms of data and available models (Moss et al., 2010b). In this light this discrepancy makes sense. It is in noteworthy that the least well understood regions are also the regions with the largest potential impact on the future of the transportation system.

4.1.3 MODEWISE PROJECTIONS

4.1.3.1 HIGH SPEED MODES

Projections indicate that HSM will for all models be the dominant mode of transportation worldwide near the end of the century. This view is supported by some literature (IETA, 2007; Schafer, 1998) but disputed by other literature (Kyle & Kim, 2011). Others still report that the TD for HSM in China will be the largest mode as early as 2040 (Zhang et al., 2007), commercial parties forecast an yearly growth of 7% (Airbus, 2012). Again more accurate data from developing region would be required to confirm either claim.

No matter the details of the projections taken, the lack of low CO_2 alternatives to aircraft (Bastien Girod, Van Vuuren, & Hertwich, 2013; J. Lee & Mo, 2011; Westenberger, 2008) means this mode will be of great impact to the climate in the future. An alternative high speed train, ignored in the present investigation due to its small relevance in most regions. This already makes up to 20% of TD in Japan (Schafer, Heywood, Jacoby, 2010). An extended investigation into this mode would therefore be needed. Lastly aggregated approach to HSM might be overly simplistic with large variations over flight distance (Njegovan, 2006) for price, and between developing and developed (IETA, 2007) for income. For example, the income elasticity of a domestic route can be as low as 0.8, whereas for an intercontinental flight between developing countries it can be 3.6 (Njegovan, 2006).

4.1.3.2 LIGHT DUTY VEHICLES

Modeling of LDV gives a small, steady growth in TD for most regions and models. Except for the POLES model, which shows a strong saturation effect, then declining TD for the industrialized

region. For the developing region it shows a strong increase in both equipment rate and TD. The latter behavior is actually quite realistic, due to a saturation of the equipment rate of LDV in a given society, as presented in (Dargay et al., 2007). It has to be noted that the same author (Dargay et al., 2007) predicts a much stronger increase in the equipment rate than any of the models at present, equipping virtually every adult on the planet with a car by 2100.

Figure 115 presents the saturation of the income elasticity that is a consequence of saturation of the number of cars (Dargay et al., 2007). The elasticity first increases when income is high enough to buy a car, then decreases when saturation is reached. Notice that no country has at yet reached the saturation level of cars yet (~90% equipment rate) and this is therefore a somewhat hypothetical framework. Another issue with LDV that rather concerns the developing region is the mass adoption of motorcycles, which increases and then decreases with economic development (Kitous, 2010a). Therefore an increase in mobility can occur much earlier than predicted based on income simply through adoption of cars.

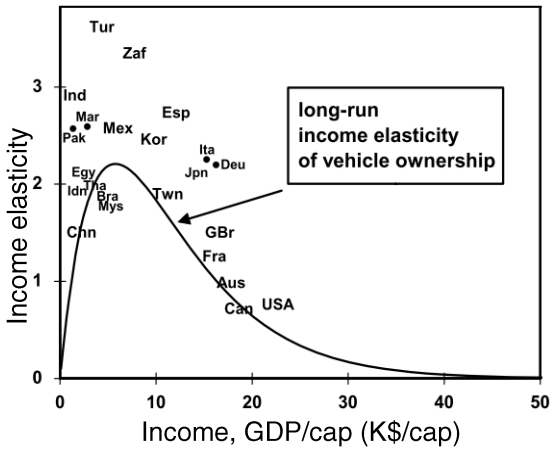


Figure 115: Income elasticity ϵ of LDV as a function of income, there is a clear saturation effect

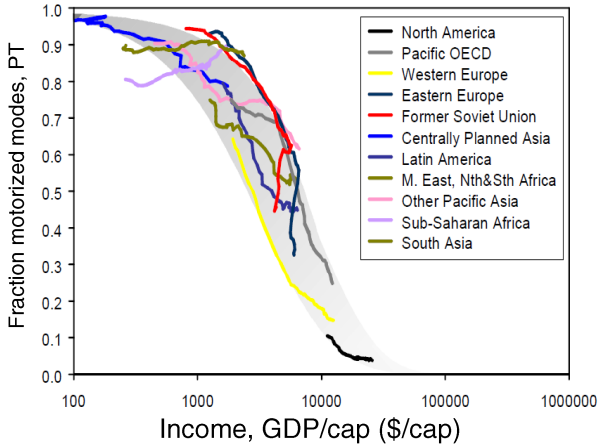


Figure 116: Development of public transport PT with economic growth (Schäfer, 2005)

4.1.3.3 PUBLIC TRANSPORTATION

The development of TD PT differs more than the other modes, between the models. TD for LDV is smaller than LDV for GCAM and TIMER-travel, as large as LDV for SIMPLE, and depending on the region, large or small for POLES. The issue with PT as an elasticity is displayed in Figure 116, with increasing income TD for PT increases, with more income growth TD increases for all modes, hence the TD for PT will then decrease. Therefore initially there is a strong positive income elasticity ϵ that later changes sign into a negative ϵ with respect to TD. This development for PT is in various stages across the world, where it is receding in some regions and upcoming in others. The negative ϵ is actually found in all literature sources except the POLES model literature, (Kitous, 2010b). A positive ϵ is used in both SIMPLE and POLES. PT is the smallest motorized mode in terms of TD, however further investigation of this mode would benefit at least the POLES model, and might make the other models more coherent.

4.1.3.4 NON-MOTORIZED

NM is not used in any comparison, they don't exist in every model investigated and aren't considered contributors to energy use and CO_2 emissions. Additionally there is not as much accurate data about their usage worldwide or their energy and emission variables, which are considered quite negligible (Wright & Fulton, 2005). They might be relevant if one uses detailed LCA analysis, because humans need food. They are also a significant time sink for the developing region which might impede economic growth. Therefore the assumptions of a time use equal to the TTB minus the time use for motorized might be oversimplified. Data is very sparse on elasticities and TD for this mode. This might improve with new investigations thatuse modern GPS technologies on individuals (Stopher & Zhang, 2011).

4.2 IMPLICATIONS

The implications are split up in the same way as the research questions, first discussion of the models, then the historic validity and lastly the probabilistic projections.

4.2.1 IMPLICATIONS FOR MODELERS

First and foremost openness, the availability of good model descriptions and access to data proved essential to successful model replication. Single publications proved impossible in recreating models due to errors in notation, ambiguity and simply because the models were too complex. Therefore cooperation with authors proved essential to the development of a successful model in this framework. If one thing could improve the work for modelers it is openness of source codes, data and technical description.

The modeling approaches investigated were the TIMER-travel, GCAM, POLES and the SIMPLE. The first two models rely on a competition between travel modes and the latter depends solely on the elasticities of individual modes. The mathematical notation of the GCAM model proved straightforward and very robust, few assumptions were made. The TIMER-travel model was more complicated to create because of several variables not mentioned in the publications and unstable regimes. The POLES model proved difficult to create due to ambiguous notations, over fitting of the model and a lack of calibrated independent variables (instead preferring regional variables). Additionally unstable regimes existed, some variable signs were inverted compared to literature and the use of saturation levels left room for interpretation. The SIMPLE model is very straightforward and simply requires a literature investigation for the correct starting values.

Some assumptions made by the model formulation proved to be successful, as distilled for all the sections above, others did not. TTB proves to be a robust assumption, although the time weight definition by TIMER-travel leads to some unstable regimes. TMB is a good predictor of a good historic fit, but possibly in a looser range then defined by Schafer (Schafer, Heywood, Jacoby, 2010). A regional correction might be needed for the TMB as used in TIMER-travel. Adherence of TU and TTB is also found to be a good indicator of a good historic fit. Therefore boundaries imposed on both TU and TTC could improve models, they could also be used as error indicators instead of the fit on historic TD data.

The two best fitting models, which the largest number of valid regions, timer-TRAVEL and GCAM could be guidelines to good modelling practice. Both have competition between modes (e.g. MNL, Logit), which divide up a summated TD. There is also the usage of some sort of income and door to door speed DDS feedback on the travel price, as defined by the comfort factor in TIMER-travel or the wage rate/DDS by GCAM. They both forecast very similar price developments. It also appears that fewer regional variables, or rather regional calibration variables, are also a good predictor of modeling performance worldwide, as demonstrated by the only two models with this feature, GCAM and TIMER-travel.

The two worst fitting models, witch less than half the regions valid, POLES and GCAM could be guidelines in what to avoid. The use of elasticities fixed in time, for each region. Additionally the use of saturation levels leads to a less robust level, and the use of a time trend also leads to

instabilities in the model. Lastly price elasticities are difficult to calculate from data, the investigation did not succeed in doing this unambiguously even during attempts with econometric methods. They are also hard to put in context when compared to literature, the literature definition of price will differ from that of the model. It can for example be fuel price, but also travel price or total cost of ownership, which comes with different assumptions on, for example, life time by different authors.

It is reasonable to think convergence are preferences to industrialized region standards plays a role in future developments, as already implemented in GCAM, a recent version of TIMER-travel and POLES, but not in SIMPLE. Literature does state convergence of energy patterns, but doesn't specifically mention transportation (Mielnik & Goldemberg, 2000). Regional differences might mean the present method is oversimplified. Nonetheless it can be shown that if saturations play a large role focus should be on low CO_2 LDV development rather than HSM, because HSM will then not be the largest mode of transport by TD in 2100.

4.2.2 IMPLICATIONS OF HISTORIC FITS

Moving on to the second part of the research question, the historic fit. All models prove they can forecast historic data for almost all regions given ten years of calibration and 35 years to forecast. This is at least when comparing the graphs of the different models visually, neglecting the error measure. The distributions of errors, TDfit, produced by the Monte Carlo simulations shows that what all models have in common is a poor historic fit for the CPA region (China). It is said that error in any model can come from three sources: the data, the model and the modeler (Ruijven et al., 2009). Assuming the latter untrue this leads to the strong suspicion that this region suffers from poor data quality. It has to be noted that South Asia and other Pacific Asia might have worse data than other regions, but this does not seem to affect the overall fit. It is, however, ironic that these regions are also the ones with the largest projected growth in TD.

Regarding historic TD the GCAM and TIMER-travel model fit well to all regions except CPA. The SIMPLE model fits poorly to developing region, and the POLES model only has an acceptably small TDfit for industrialized nations. The adherence of projections to the TTB and TMB have also been tested, and it proves the regions that these boundaries are adhered to are the same ones that adhere to a good TD fit.

4.2.3 IMPLICATIONS OF FUTURE PROJECTIONS

This section discusses the third part of the investigation, the probabilistic projections that have been made. The reader is invited to investigate the full page world map of Figure 106 or Figure 114. It can be seen that although TD for the industrialized and developing region are in 2005 nearly equal, by 2100 the TD in the developing region is likely to have increased to 8-16 times 2005 levels. With the TD in the industrialized region at most increasing 2-4 times. The concern regarding CO_2 emissions and should hence be moved to the developing region. The TD in the developing region is expected to increase at least 10-fold and at most 20-fold from 2005 levels by 2100. The concern is that the CO_2 emissions that follow this increase will become a major issue in the context of global warming.

Dissecting the graphs into individual contributions for each mode shows that by 2060 HSM will be the dominant form of transportation worldwide, 2 - 3 times the TD for LDV, or at equal in TD to LDV in the converging preference models. HSM will have the number of TD increased to roughly 10 to 20 times the level of 2005. PT will not play as much of a role in the future as either modes, TD for PT is at most half as much as TD for LDV (GCAM, TIMER-travel). This is in line with PT declining with income growth (Litman, 2011). Notice that in terms of the time spent in vehicles PT is actually significant and HSM much smaller.

4.3 LIMITATIONS

This section will first briefly discuss the limitations of the comparative method developed, discuss the results of the projections and lastly discuss implications for modeling methods.

Four models were built, compared and analyzed. The approach in combining models of different types, with different core concepts and different datasets into a single framework appears to be a good approach in comparing models in a very controlled way, it also appears to be relatively novel. Given enough time, computational power and the right technical documentation model, understanding can be improved using these methods. Although the method clearly works, it would benefit from having access to original results, input data and direct interfacing with the authors at a late stage in the investigation. Too much exposure might however affect the objectivity the investigator has when looking at the models afresh, so a balance should be struck.

The methods used in the present investigation are attempts at keeping the models and comparison unbiased. Assumptions and datasets were based on general literature rather than solely author published values, calibration was performed by automated methods and the entire variable space was mapped before ranges for the Monte Carlo simulations were selected. Additionally projections were only analyzed after the models had been developed and historically tested. Lastly generating probabilistic projections is computationally intensive and hence tweaking post-hoc is difficult and ranges had to be selected very carefully.

Despite the precautions some choices had to be made, such as the base dataset. Most of the dataset used came from TIMER (De Vries et al., 2001), TIMER-travel (Bastien Girod et al., 2012) and Schafer (Schafer, Heywood, Jacoby, 2010) these datasets might benefit the TIMER-travel model that has been designed on them. Additionally some variables such as price elasticities, are notoriously hard to derive from data unambiguously, even with econometric methods. The elasticities are also often regionally and mode dependent. Another example is the ambiguity in the POLES model, where time trend, equipment rate, ϵ and σ all work on a single dimension, TD per mode. This leads to over fitting, and hence a subjective selection of variable values. The independent variables here were derives almost exclusively from western literature that might be biased toward the Industrialized region.

It has proven very difficult to replicate a model solely based on published work and data, although all authors in the present investigation were more than willing to cooperate. The latter fact makes for an interesting debate on whether published results of this complex kind alone are transparent enough ensure the boundary condition of replication are met. This is vital to the objectivity of science as defended by David Goodstein in his book "On Fact and Fraud" (Goodstein, 2010). Time efficient analysis requires flexible tools for among others the optimizations, Monte Carlo simulations, post processing, analysis and visualization tools. All these tools had to be developed from scratch for the present investigation and a future investigation of the kind could be simplified significantly if these tools were available. The Matlab toolbox used for this investigation will be available open source from https://globaltravelmodels.codeplex.com.

4.4 FUTURE DIRECTIONS

This section will first discuss the probabilistic projections, next the analysis method, then new research and lastly suggestions for improved models

The probabilistic projections made regarding TD have proven comparable to other literature and provided an insight into the future of passenger transportation and hence global climate change. The probabilistic projections created can provide insight with regards to the development of range in the models. The modifications suggested to the models here should be investigated further, and lessons learned could be applied. For example, the converging of preferences added to the models here could be investigated further, it could additionally be determined if such effects are justified, or rather the unsaturated models should be used.

The analysis method of this publication is novel and has proven useful in the investigation of several different types of models. It could be applied to a broad range of modeling fields, as long as models are small enough for relatively fast computations. The modeling and analysis toolbox created in Matlab could be useful in speeding up future model investigations. Applications of this type of research could be to improve understanding of a given set of models, to create fairly objective probabilistic projections and to be able to compare how realistic different modeling approaches are in predicting future events. The toolbox and the methods can be used in future investigation and are freely downloadable

In future research several improvements can be suggested, including the use of an independent dataset and additional technical documentation for the replication of the models. The more sources the lower the ambiguity, as each source in detail has some ambiguities or errors left by the best authors. In regard to the field of transportation modeling, accuracy would improve with a better dataset of the developing regions. A more detailed investigation of vehicle types and fuel types might be considered. A deeper investigation of the CO_2 and energy scenarios could be combined with the TD results already presented to infer direct climate implications, in the present investigation this was not a main focus. Lastly the inclusion of more uncertainties such as income or population growth could provide more insights into the range of the forecasts (Energy systems division, 2009). This can be done either by including the uncertainties as ranges, as has been done with other variables, or into scenarios, such as done by other authors (Van Vuuren, 2007). The latter seems more sensible.

If the lessons of the present investigations are combined it might be possible to produce a hybrid model. That takes the best of each model and implements it into a single model with a single prediction. Such a model should contain the TMB and TTB restrictions, TIMER-travel therefore seems like a good starting point. It should also contain more detailed assumptions about TD as a function of price and income. Additionally65 globally defined variables with regional calibration variables, mode competition and convergence of some preference factor could be added. It could even contain cross-modal effects, such as those between PT and LDV. Additionally it could contain more detailed modes such as the distinction between motorcycle and car for Asian regions. Lastly macro-economic data such as wage rates and equipment rates could improve the model by adding more macro-economic details, without losing the aggregated scoping.

5 CONCLUSION

It can be concluded that a method for investigating a sample of models with different methodologies, datasets and assumptions in a detailed and objective manner has been developed. It was successfully applied to four different models and shed light on the mechanisms that made them function, compared them a critically analyzed their ability to represent historic data. It produced probabilistic projections that cover the range of published non-probabilistic projections. The conclusion will be split up according to the three pillars of the research question, first the modeling formulations, second the historic fit and third the projections.

Firstly we can say that the modeling formulations compared were quite different and led to varying results for the models. The models with mode competition and a component that covers the value of time in their mode pricing had better historic fits than the two models which are based on modal elasticities. Between them these models produce similar predictions at the modal, regional and aggregated regional level. The SIMPLE and POLES models consisted of solely elasticities and the latter of elasticities with time trends, saturations and equipment rate development. The latter was valid in industrialized nations but had stability issues near saturation limits for High Speed Modes HSM and Light Duty Vehicles LDV. From these facts one can conclude that the fixed elasticity models as described in the present investigation lack the dynamics required for a global long term travel model, they cannot recreate the historic datasets accurately, and did not meet empirical facts such as the Travel Time Budget TTB and Travel Money Budget TMB. This does not mean all elasticity definition are wrong, but it rather makes the case that they are oversimplified and require further refinement.

Secondly the historic fit, which can be split up in the fit during the validation phase and the Monte Carlo phase. During the former the models were asked to forecast the TD 1981-2005, based on calibration from 1971-1980. With the variables rather broadly defined each model could fit to almost every region. Which means given enough freedom each model could be considered historically sound. However, with the Monte Carlo experiments and stricter limits on variables based on the methods described. The fit for Centrally Planned Asia (China) was poor for all models, and was insufficient for the developing region for SIMPLE and POLES. Further investigation into the adherence time use and total travel cost to realistic TTB and TMB values yielded more reliable results and led to the rejection of the POLES model for non-industrialized region. These two criteria were used in selecting the valid model regions and proved to be reliable indicators for a good model fit, providing similar results as the error on the fit. They are hence important constraints for a model to follow.

Thirdly the projections, the projections are split up into saturated and non-saturated. For the both sets the future TD give a large increase in the TD, more than 90% of this will be due to developing region, which only make up 50% in 2005. For the original models without convergence the majority mode of transportation will be high speed modes (beyond 2060). For models with convergence of preferences to industrialized standards LDV and HSM converge to a similar level by 2100. The increase is projected to be 8-16 times 2005 levels of TD for the world compared to a 2-4 fold increase in industrialized nations. The amount of time spent per mode is in all models rather related to LDV and PT than HSM, because HSM covers 8 times as many kilometers in the same time. A detour to the implications of the rapid growth of air travel predicted by all models

was made. If air travel doesn't become less emission intensive it will make the RCP scenarios, even in the highest emission scenario, difficult to realize. Technological feasibility with current technology is therefore low and other channels should be considered for lowering these emissions.

6 BIBLIOGRAPHY

A. Kitous. (2000). Poles prospective outlook on long-term energy systems. Grenoble.

Airbus. (2012). *Global Market Forcast*.

Ajanovic, A., Dahl, C., & Schipper, L. (2012). Modelling Transport (Energy) Demand and Policies – an introduction. *Energy Policy*, *41*, iii–xiv. doi:10.1016/j.enpol.2011.12.033

Avineri, E. (2009). The effect of household size and structure on travel / activity patterns : Exploration of household travel time budget. The 12th International Conference on International Association on Travel Behavior Research. Jaipur, Rajasthan, India. December 13-18, 2009.

Botterweg, P. (1995). The user's influence on model calibration results: an example of the model SOIL, independently calibrated by two users. *Ecological Modelling*, *81*(1-3), 71–81. doi:10.1016/0304-3800(94)00161-A

Bouwman, A. F., Kram, T., & Goldewijk, K. K. (2006). *Integrated modelling of global environmental change*. PBL.

Chamon, M. Mauro, P. Okawa, Y. (2008). Mass car ownership in the emerging market giants. *Economic Policy*, (April).

Chen, C., & Mokhtarian, P. (2008). A Review and Discussion of the Literature on Travel Time and Money Expenditures. (I. of T. Studies, Ed.)*eScholarShip California*.

Crick, M. J., Hill, M. D., & Charles, D. (1987). The Role of Sensitivity Analysis in Assessing Uncertainty (pp. 1–258). Proceedings of an NEA Workshop on Uncertainty Analysis for Performance Assessments of Radioactive Waste Disposal Systems.

Currie, G., & Phung, J. (2006). Exploring the impacts of fuel price increases on public transport use in Melbourne. Currie, G., & Phung, J. (2006). Exploring the impacts of fuel price increases on public transport use in Melbourne.

Dahl. (2011). Dahl Energy Demand Database. *Modelling transport energy demand*. Retrieved from http://dahl.mines.edu/courses/dahl/dedd/

Dargay, J., Gately, D., & Sommer, M. (2007). *Vehicle Ownership and Income Growth , Worldwide : 1960-2030* (pp. 1–32). Leeds.

Davies, S. . (2004). The Great Horse-Manure Crisis of 1894. *The Freeman*.

De Haan, P. (2009). Applied comprehensive NO 2 and particulate matter dispersion modelling for Switzerland. *Int. J. Environment and Pollution*, *36*(2), 204–223.

De Vries, B., Van Vuuren, D. P., Den Elzen, M., & Janssen, M. (2001). *The Targets IMage Energy Regional (TIMER) Model - Technical documentation* (2001st ed.). PBL Netherlands Environmental Assessment Agency.

Deetman, S., Hof, A. F., Pfluger, B., Van Vuuren, D. P., Girod, B., & Van Ruijven, B. J. (2013). Deep greenhouse gas emission reductions in Europe: Exploring different options. *Energy Policy*, *55*, 152–164. doi:10.1016/j.enpol.2012.11.047

Di Lorenzo, G., Pilidis, P., Witton, J., & Probert, D. (2012). Monte-Carlo simulation of investment integrity and value for power-plants with carbon-capture. *Applied Energy*, *98*, 467–478. doi:10.1016/j.apenergy.2012.04.010

EAR. (2012). *National Energy Review 2012*.

Edenhofer, O., Ramón Pichs-Madruga, Y. S., Seyboth, K., Patrick Matschoss, S. K., Zwickel, T., Eickemeier, P., Hansen, G., et al. (2011). *Special Report on Renewable Energy Sources and Climate Change Mitigation*. IPCC publishing.

Edmonds, J. A., Wise, M. A., & MacCracken, C. N. (1994). *Advanced Energy Technologies and Climate Change: An Analysis Using the Global Change Assessment Model (GCAM)*.

Energy systems division. (2009). *Multi-Path Transportation Futures Study : Vehicle Characterization and Scenario Analyses*.

Federal Aviation Authority. (2012). *FAA Aerospace Forecast Fiscal Years 2012-2031* (pp. 1–34).

Fulton, L., & Eads, G. (2004). *IEA/SMP Model documentation and reference case projection*. *Open publication*.

Girod, B., Van Vuuren, D., & De Vries, B. (2013). Influence of Travel Behavior on Global CO 2 Emissions. *Transportation Research Part A: Policy and Practice*, *50*, 183–197.

Girod, Bastien, Deetman, S., & Van Vuuren, D. P. (2011). Global transportation and energy use in the next decades. *Model description and publication*.

Girod, Bastien, Van Vuuren, D. P. \, Grahn, M., Kitous, A., & Kyle, P. (2013). Climate Impact of Transportation A Model Comparison. *Climatic Change*, *118*(3-4), 595–608.

Girod, Bastien, Van Vuuren, D. P., & De Vries, B. (2013). Influence of travel behavior on global CO2 emissions. *Transportation Research Part A: Policy and Practice*, *50*, 183–197. doi:10.1016/j.tra.2013.01.046

Girod, Bastien, Van Vuuren, D. P., & Deetman, S. (2012). Global travel within the 2°C climate target. *Energy Policy*, *45*, 152–166. doi:10.1016/j.enpol.2012.02.008

Girod, Bastien, Van Vuuren, D. P., Grahn, M., Kitous, A., Kim, S. H., & Kyle, P. (2013). Climate impact of transportation A model comparison. *Climatic Change*, (3). doi:10.1007/s10584-012-0663-6

Girod, Bastien, Van Vuuren, D. P., & Hertwich, E. G. (2013). Corrigendum: Global climate targets and future consumption level: an evaluation of the required GHG intensity. *Environmental Research Letters*, *8*(2), 029501. doi:10.1088/1748-9326/8/2/029501

Goodstein, D. (2010). *On fact and Fraud*. Princeton University Press.

Graham, D. J., & Glaister, S. (2004). Road Traffic Demand Elasticity Estimates: A Review. *Transport Reviews*, *24*(3), 261–274. doi:10.1080/0144164032000101193

Grahn, M. (2007). *Cost effective fuel choices in the transportation sector under different international climate regimes - results from a regionalized version of the global energy transition model, GET-R. Grahn, Maria, Christian Azar, and Kristian Lindgren. "Cost effective fuel choices in the transportation sector under different international climate regimes–results from a regionalized version of the global energy transition model, GET-R." Proceedings Boo*. Gothenberg, Sweden: Chalmers university press.

Hamby, D. (1994). A review of techniques for parameter sensitivity analysis of environmental models. *Environmental Monitoring and Assessment*, *32*(2), 135–154.

Hedenus, F., Karlsson, S., Azar, C., & Sprei, F. (2010). Cost-effective energy carriers for transport – The role of the energy supply system in a carbon-constrained world. *International Journal of Hydrogen Energy*, *35*(10), 4638–4651. doi:10.1016/j.ijhydene.2010.02.064

Hepburn, C., & Müller, B. (2010). International Air Travel and Greenhouse Gas Emissions: A Proposal for an Adaptation Levy1. *World Economy*, *33*(6), 830–849. doi:10.1111/j.1467-9701.2010.01287.x

IATA Consulting. (2011). *Air Travel Demand Economic Briefing*.

IETA. (2007). *Estimating Air Travel Demand Elasticities Final Report*.

International Energy Agency. (2007). *Climate Policy Uncertainty and Investment Risk*.

International Energy Agency. (2012). *Key world energy statistics*.

IPCC. (2007). *IPCC fourth assessment report*.

Katz, R. (2002). Techniques for estimating uncertainty in climate change scenarios and impact studies. *Climate Research*, *20*, 167–185.

Keynes, J. M. (2007). *Working Time Around the World*. International Labour Office.

Khandelwal, B. (2011). Study on Nuclear Energy Powered Novel Air Transport Model and its Feasibility study. *Joint propulsion conference*.

Kim, S. H., Edmonds, J., Lurz, J., Smith, S. J., & Wise, M. (2006). The o bj ECTS Framework for integrated Assessment : Hybrid Modeling of Transportation. *Energy Journal*, *Special is*(Special Issue# 2), 63–91. Retrieved from http://ideas.repec.org/a/aen/journl/2006se_jaccard-a04.html

Kitous, A. (2010a). *Transport in POLES - manual*.

Kitous, A. (2010b). Prospective Outlook on Long-Term Energy Systems - Manual.

Klepper, O. (1997). Multivariate aspects of model uncertainty analysis: tools for sensitivity analysis and calibration. *Ecological Modelling*, *101*(1), 1–13. doi:10.1016/S0304-3800(96)01922-9

Kupiainen, K., & Klimont, Z. (2007). Primary emissions of fine carbonaceous particles in Europe. *Amospheric Environment*, *41*, 2156–2170. doi:10.1016/j.atmosenv.2006.10.066

Kyle, P., & Kim, S. H. (2011). Long-term implications of alternative light-duty vehicle technologies for global greenhouse gas emissions and primary energy demands. *Energy Policy*, *39*(5), 3012–3024. doi:10.1016/j.enpol.2011.03.016

Laborsta. (2012). Laborsta Labor database. Retrieved from http://laborsta.ilo.org

Lee, J. J. (2000). *Historical and Future Trends in Aircraft Performance , Cost , and Emissions by*. Massachusetts Institute of Technology.

Lee, J., & Mo, J. (2011). Analysis of technological innovation and environmental performance improvement in aviation sector. *International journal of environmental research and public health*, *8*(9), 3777–95. doi:10.3390/ijerph8093777

Litman, T. (2011). *Transportation Elasticities: How Prices and Other Factors Affect Travel Behavior*.

Litman, T. (2012a). *Understanding Transport Demands and Elasticities How Prices and Other Factors Affect Travel Behavior*.

Litman, T. (2012b). Transit Price Elasticities and Cross-Elasticities. *Journal of Public Transportation*, *7*(2), 37–58.

Litman, T. (2012c). *Understanding Transport Demands and Elasticities How Prices and Other Factors Affect Travel Behavior*.

Louviere, J. J., & Woodworth, G. (1983). Choice Allocation Consumer Experiments : An Approach Aggregate Data, *20*(4), 350–367.

Marler, R. T., & Arora, J. S. (2004). Survey of multi-objective optimization methods for engineering. *Structural and Multidisciplinary Optimization*, *26*(6), 369–395. doi:10.1007/s00158-003-0368-6

Mathworks. (2012). Matlab 2012a.

Mielnik, O., & Goldemberg, J. (2000). Converging to a common pattern of energy use in developing and industrialized countries. *Energy Policy*, *28*(8), 503–508. doi:10.1016/S0301-4215(00)00015-X

MIT Energy Initiative. (2008). *On the Road in 2035*. MIT Press.

Modelling, E., & Bilthoven, B. A. (1995). Ecologigal modelling, *83*, 55–66.

Moss, R. H., Edmonds, J. a, Hibbard, K. a, Manning, M. R., Rose, S. K., Van Vuuren, D. P., Carter, T. R., et al. (2010a). The next generation of scenarios for climate change research and assessment. *Nature*, *463*(7282), 747–56. doi:10.1038/nature08823

Moss, R. H., Edmonds, J. a, Hibbard, K. a, Manning, M. R., Rose, S. K., Van Vuuren, D. P., Carter, T. R., et al. (2010b). The next generation of scenarios for climate change research and assessment. *Nature*, *463*(7282), 747–56. doi:10.1038/nature08823

New, M., & Hulme, M. (2000). Representing uncertainty in climate change scenarios : a Monte-Carlo approach. *Integrated Assessment*, *1*, 203–213.

Njegovan, N. (2006). Elasticities of demand for leisure air travel: A system modelling approach. *Journal of Air Transport Management*, *12*(1), 33–39. doi:10.1016/j.jairtraman.2005.09.003

Paulley, N., Balcombe, R., Mackett, R., Titheridge, H., Preston, J., Wardman, M., Shires, J., et al. (2006). The demand for public transport: The effects of fares, quality of service, income and car ownership. *Transport Policy*, *13*(4), 295–306. doi:10.1016/j.tranpol.2005.12.004

Rodney, E., Falvey, E., & Gemmell, N. (1996). Are services income-elastic? Some new evidence. *Review of Income and Wealth*, *42*(3).

Ruijven, B., Sluijs, J. P., Van Vuuren, D. P., Janssen, P., Heuberger, P. S. C., & Vries, B. (2009). Uncertainty from Model Calibration: Applying a New Method to Transport Energy Demand Modelling. *Environmental Modeling & Assessment*, *15*(3), 175–188. doi:10.1007/s10666-009-9200-z

Schafer, A. (1998). The global demand for motorized mobility. *Transportation Research Part A: Policy and Practice*, *32*(6), 455–477.

Schäfer, A. (2005). *Transportation , Energy , and Technology in the 21st Century*.

Schäfer, A., & Jacoby, H. D. (2005). Technology detail in a multisector CGE model: transport under climate policy. *Energy Economics*, *27*(1), 1–24. doi:10.1016/j.eneco.2004.10.005

Schafer, A., & Victor, D. G. (2000). The future mobility of the world population. *Transportation Research Part A: Policy and Practice*, *34*(3), 171–205. doi:10.1016/S0965-8564(98)00071-8

Schafer, Heywood, Jacoby, W. (2010). *Transportation in a climate constrained world* (First edit.). The MIT Press.

Sofer, I. G. S. G. N. A. (2009). *Linear and nonlinear optimization* (2 edition .). Society for Industrial Mathematics.

Sparke, L. (n.d.). *Reducing Emissions Associated with Electric Vehicles*. Melbourne, Australia.

Stopher, P., & Zhang, Y. (2011). Travel time expensitures and travel time budgets - prelimiary findings.

Storchmann, K. (2005). Long-Run Gasoline demand for passenger cars: the role of income distribution. *Energy Economics*, *27*(1), 25–58. doi:10.1016/j.eneco.2004.03.002

Suganthi, L., & Samuel, A. a. (2012). Energy models for demand forecasting—A review. *Renewable and Sustainable Energy Reviews*, *16*(2), 1223–1240. doi:10.1016/j.rser.2011.08.014

Susilo, Y., & Avineri, E. (2011). *The impacts of household structure to the individual stochastic travel and out of-home activity time budgets*.

United Nations. (2009). *World Population Projections 2009*.

Van Vuuren, D. P. (2007). *Energy systems and climate policy : Long-term scenarios for an uncertain future*. Unknown.

Westenberger, A. (2008). *LH2 as Alternative Fuel for Aeronautics – Study on Aircraft Concepts* (pp. 1–14).

Wohlgemuth, N. (1998). World transport energy demand modelling Methodology and elasticities I. *Energy Policy*, *25*(97), 1109–1119.

Wright, L., & Fulton, L. (2005). Climate Change Mitigation and Transport in Developing Nations. *Transport Reviews*, *25*(6), 691–717. doi:10.1080/01441640500360951

Yáñez, M. F., Raveau, S., & Ortúzar, J. D. D. (2010). Inclusion of latent variables in Mixed Logit models: Modelling and forecasting. *Transportation Research Part A: Policy and Practice*, *44*(9), 744–753. doi:10.1016/j.tra.2010.07.007

Zhang, S., Jiang, K., & Liu, D. (2007). Passenger transport modal split based on budgets and implication for energy consumption: Approach and application in China. *Energy Policy*, *35*(9), 4434–4443. doi:10.1016/j.enpol.2007.03.007

A. APPENDIX: ITEMS LEFT OUT OF SCOPE

Disclaimer, results in this section where left out of the main thesis because their validity cannot be guaranteed. Therefore no value can be attached to them, they are here for illustrative purposes.

A.1 RCP SCENARIO COMPARISON

A.1.1 EMISSION SCENARIOS

The authors of all the models compared in the present investigation combine their results with efficiency projections in order to attain CO_2 emissions and energy usage (Di Lorenzo, Pilidis, Witton, & Probert, 2012). In order to provide similar insight two methods are considered.

The first method uses the TIMER (2012) baseline projections for the CO_2 equivalent emissions per kilometer per travel mode. This methodology is used by the authors of the models, but is not the focus of the present investigation. Results are therefore speculative and the graphs hence produced (example in **Error! Reference source not found.**) are not used in the conclusion and moved to the appendix.

The second method is based on a recent publication into the impact of future consumption level on the climate and implications of the RCP scenarios (Bastien Girod, Van Vuuren, & Hertwich, 2013). The publication compares the projected emissions from HSM to what is predicted by the RCP scenarios. This is interesting to investigate because there is evidence that, unlike other modes of transport, it is difficult to reduce the energy intensity of aircraft to a sustainable level. In order to put this into perspective it is tested both how much CO_2 is actually emitted and what percentage of total travel could be high speed, given the CO_2 restraints.

The data for the RCP scenarios can be found in **Error! Reference source not found.** (Moss et al., 2010b). The scenarios vary in their implications and the emissions required to get to that stage. For example RCP 8.5 projects a strong increase in CO_2 concentration by the end of the century, whilst RCP 2.6 predicts very low concentrations. Given the recently published hypothesis that aircraft are unlikely to increase much in terms of efficiency as a shift in fuels is very difficult (Bastien Girod, Van Vuuren, & Hertwich, 2013), once can put the TD by aircraft in 2100 into perspective of the RCP scenarios and determine if they are likely to be met considering aircraft (HSM) alone.

A.1.2 NO CONVERGENCE

It was found in the preceding projections that the TD for aircraft, HSM, is set to increase rapidly. In order to analyze the implications of the TD for HSM the emissions are calculated using emission intensities from the year 2005. The emissions hence found are compared to the transportation emissions in the 4 RCP scenarios.

The emissions calculated in this manner are presented in Figure 117. By 2050 over 50% of RCP emissions are caused by HSM. By 2100 more emissions are caused by HSM than by RCP

transportation emissions as a whole. The percentage of the total RCP transport emissions caused by HSM are presented in Figure 118. Without any technological improvements reducing the emission intensity of HSM none of the RCP scenarios can realistically be met by 2100.

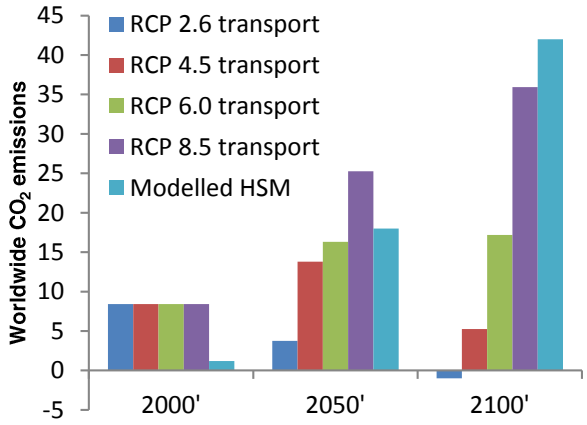


Figure 117: Average worldwide CO_2 emissions caused by HSM at the emission intensity of 2000, compared to the total CO2 emissions by transportation per RCP scenario. Original models

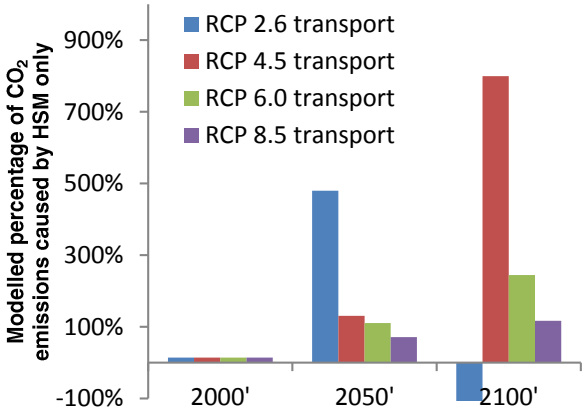


Figure 118: Percentage of expected CO_2 emissions related to transportation per RCP scenario, caused by HSM alone using emission levels of 2000. Original models

A.1.3 CONVERGENCE

It was found in the preceding projections that the TD for HSM is set to increase rapidly, be it less than with the original models. The emissions calculated in are presented in Figure 119. It can be seen that by 2050 most of the transportation emissions are caused by HSM and by 2100 more emissions are caused by HSM than three of the RCP scenarios for transportation as a whole. Only the highest emission scenario does not fully consist of emissions from HSM. The percentage of the total transport emissions caused by HSM, calculated in this manner, is presented in Figure 120. It is shown that without any technological improvements lowering the emissions intensity of HSM all but the RCP 8.5 scenario are exceeded by HSM alone by 2100.

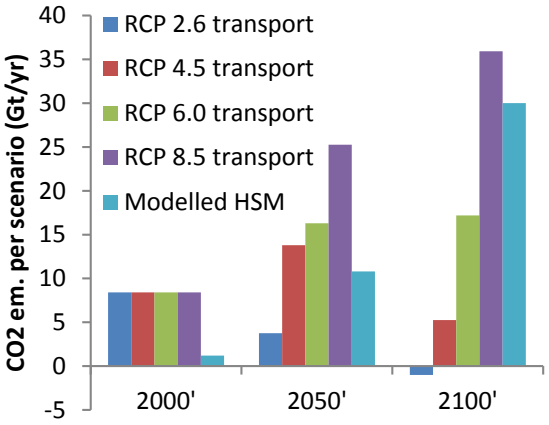


Figure 119: Average modeled worldwide CO_2 emissions caused by HSM at the emission intensity of 2000, compared to the total CO2 emissions by transportation per RCP scenario. Converging preference models

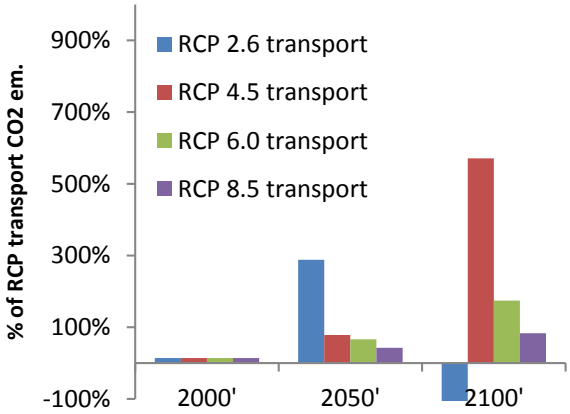


Figure 120: Percentage of expected CO_2 emissions related to transportation per RCP scenario, caused by HSM alone with convergence and emission levels of 2000. Converging preference models

A.2 CO_2 AND ENERGY PROJECTIONS

The projections in this section are no longer in the main article because their approach is rather simplistic, other methods have been applied. Nonetheless the reader might find these graphs insightful and they are hence listed here.

A.2.1 ENERGY & CO_2 WORLD SUMMARIZED

Figure 121 contains the energy use and CO_2 emission predictions worldwide. It can be seen that by far the largest consumer of energy and producer of CO_2 per capita will be the high speed mode. This is both due to a lack of CO_2 low alternatives and because of the increase in the size of this mode. On a per capita basis CO_2 emissions are in fact likely to decrease for the LDV and PT modes, whilst retaining the same level of energy use.

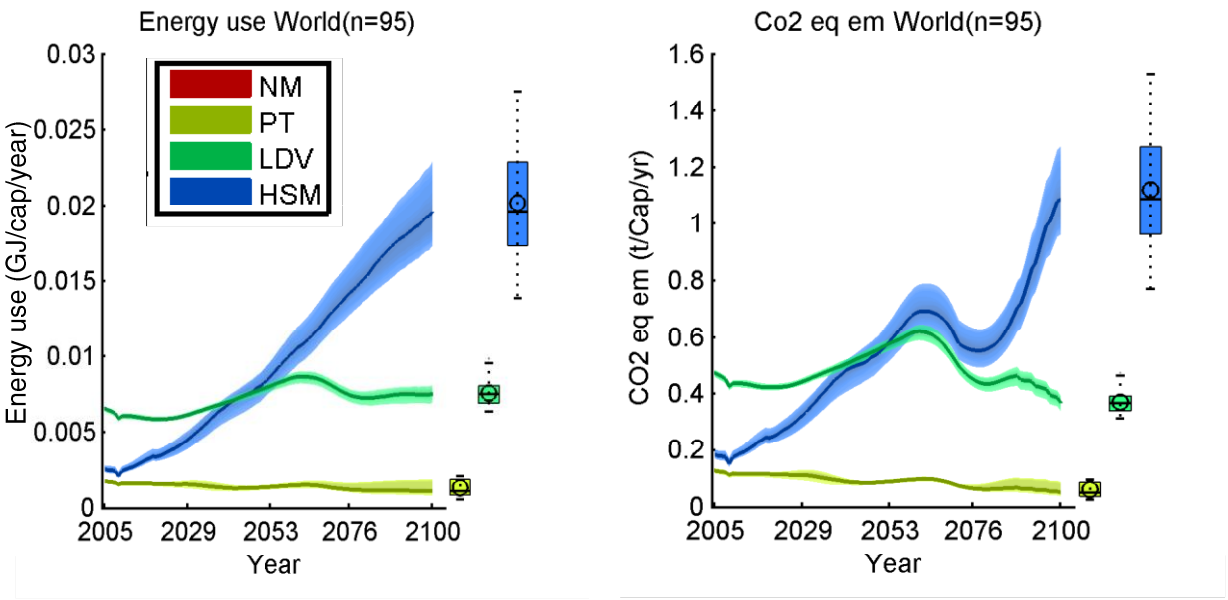


Figure 121: Energy use and CO_2 emissions per capita worldwide (except CPA) for the TIMER-travel model

The energy use and CO_2 emissions both have to HSM as their main contributor by the year 2050. Energy efficiency for LDV and PT is likely to lower the energy usage after 2050 and the same story applies for the CO_2 emissions. On a per capita basis the energy demand for HSM is likely to saturate, but not decrease and the CO_2 emissions are likely to increase over the entire period.

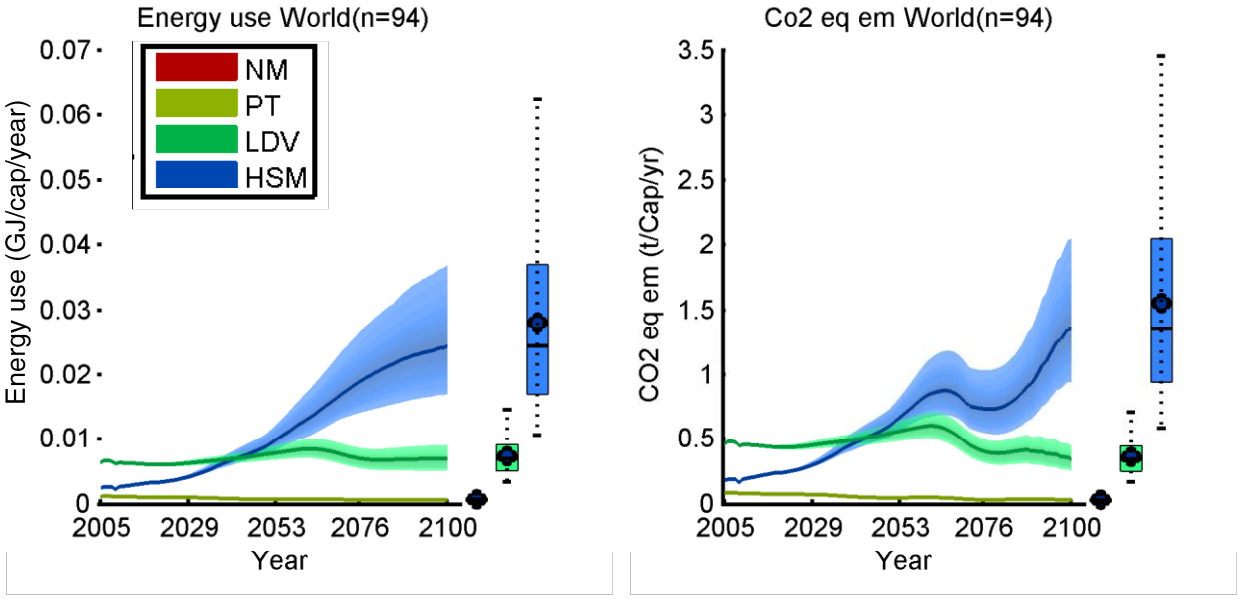


Figure 122:Energy use and CO_2 emissions worldwide for all modes in the GCAM model

Depicted in Figure 123 are the CO_2 and energy projections. The CO_2 emissions and energy use from the industrialized region shows that LDV related energy use and emissions will drop significantly over the century. PT will replace part of the lost mobility with low carbon

alternatives and HSM is likely to become the major contributor by 2060 according to most scenarios.

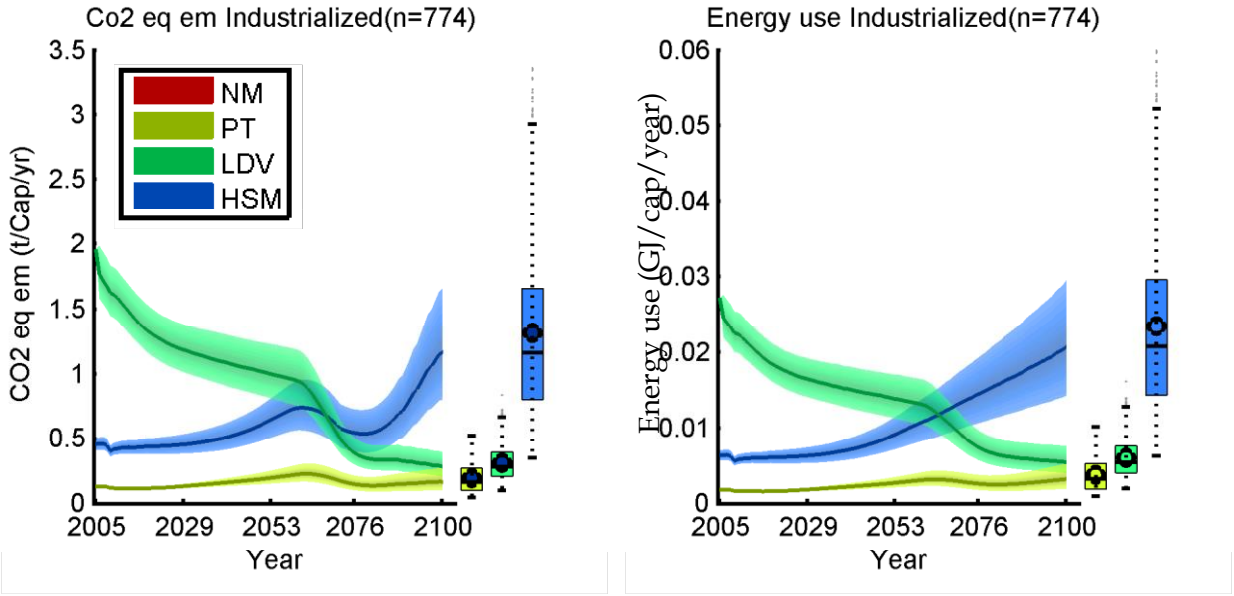


Figure 123: :Energy use and CO_2 emissions worldwide for all modes in the POLES model

Figure 124 depicts the CO_2 emissions and energy usage projections. It can be seen that HSM only slowly catches up with LDV in terms of emissions and energy use per capita. In fact it is likely not to overtake the LDV mode before 2100 in the median scenario.

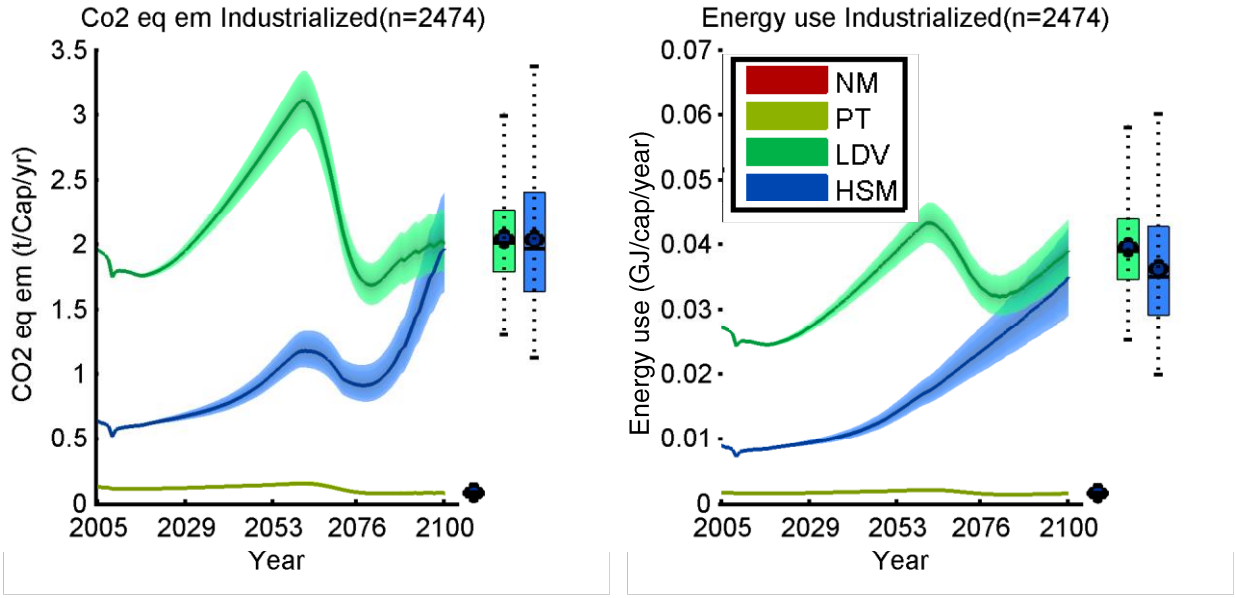


Figure 124: CO_2 and energy use projections for the SIMPLE model, per mode.

A.2.2 CO_2 AND ENERGY PER MODEL

Figure 125 depicts the CO_2 emissions trend for each of the models. It can be seen that a similar trend exists as did for the TD. The Developing region will in all scenarios become the major contributor to global CO_2 emissions by 2050. It will then continue to grow until total worldwide CO_2 emissions are between roughly two and four times what they are today. The difference between models is slightly larger than before, especially when concerned with the industrialized nations. Again, however, no significant difference between the 2100 outcomes exist at p=5% levels. This is rather due to the spread than the lack of difference between the lowest and highest values (1 and 6 Gt/yr).

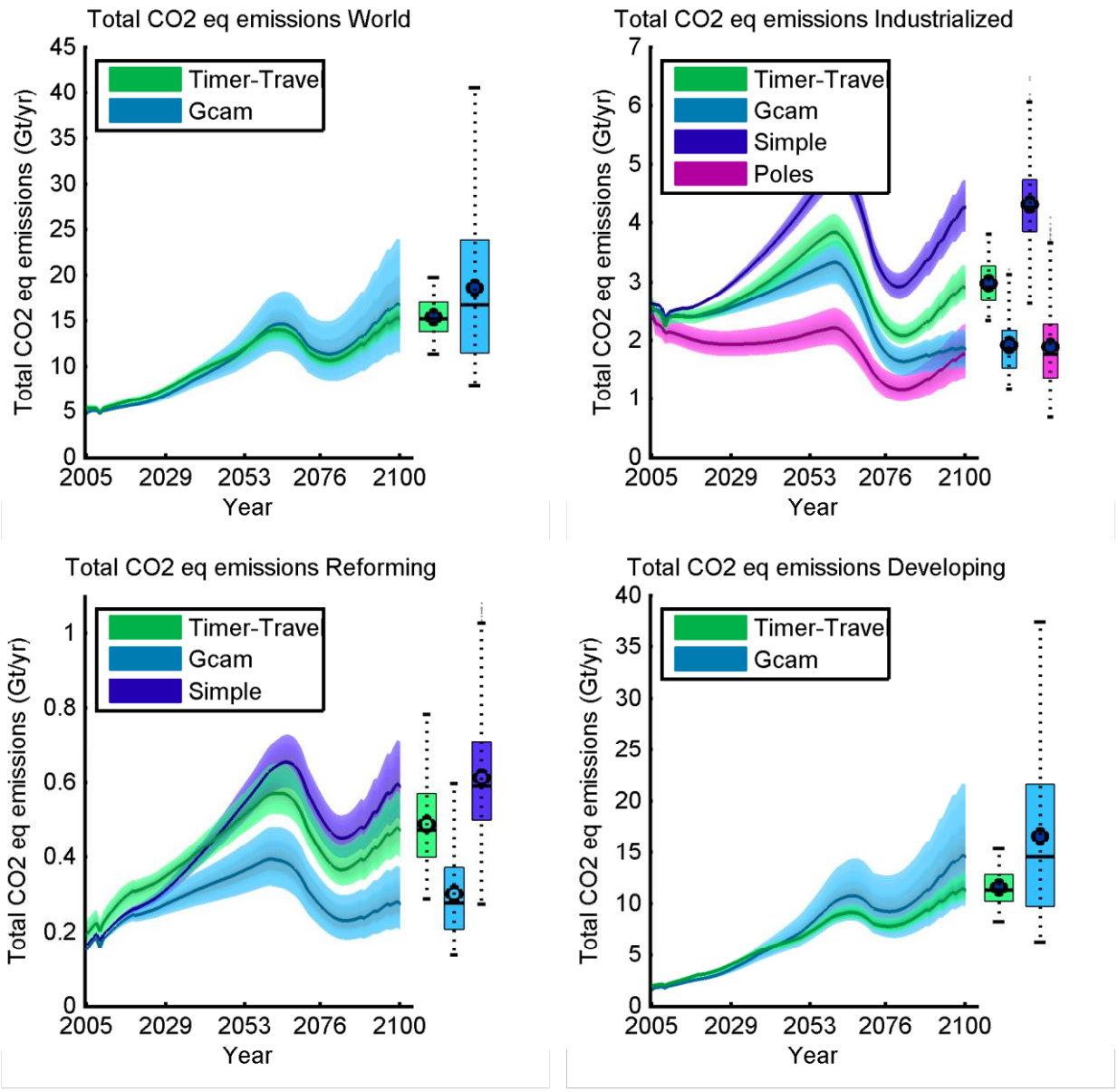


Figure 125: CO_2 emission projections per world region per model

Figure 126 presents the probabilistic projections for each economic world region. The projections are initially fairly similar to those of the CO_2 emissions, as most reductions will be due to efficiency increases and mode changes. However, as technology progresses and low CO_2 fuels become available the energy use is seen to increase much more than the CO_2 emissions.

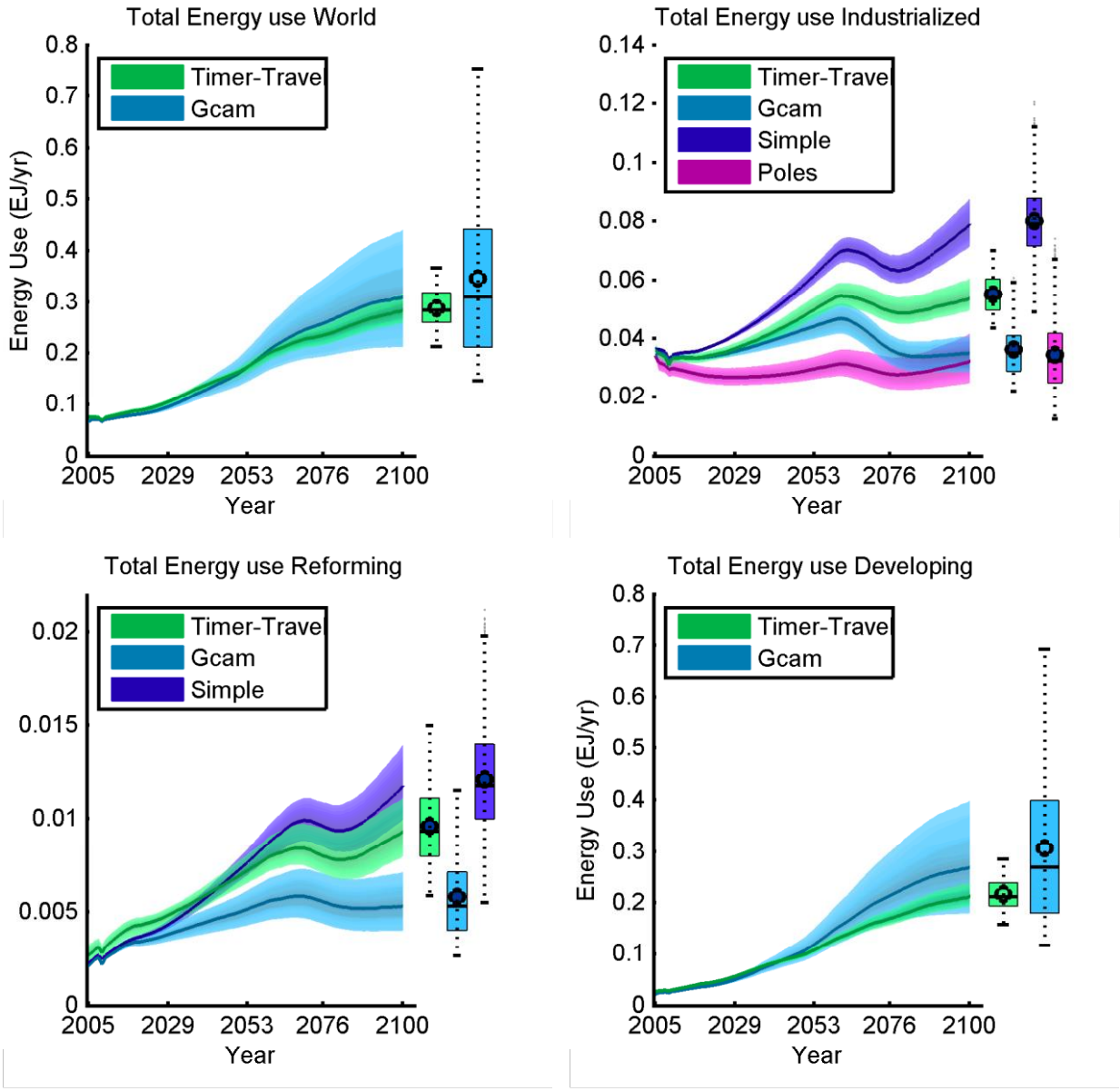


Figure 126: Energy usage projections per world region per model

A.2.3 CO_2 EMISSIONS WORLDMAP

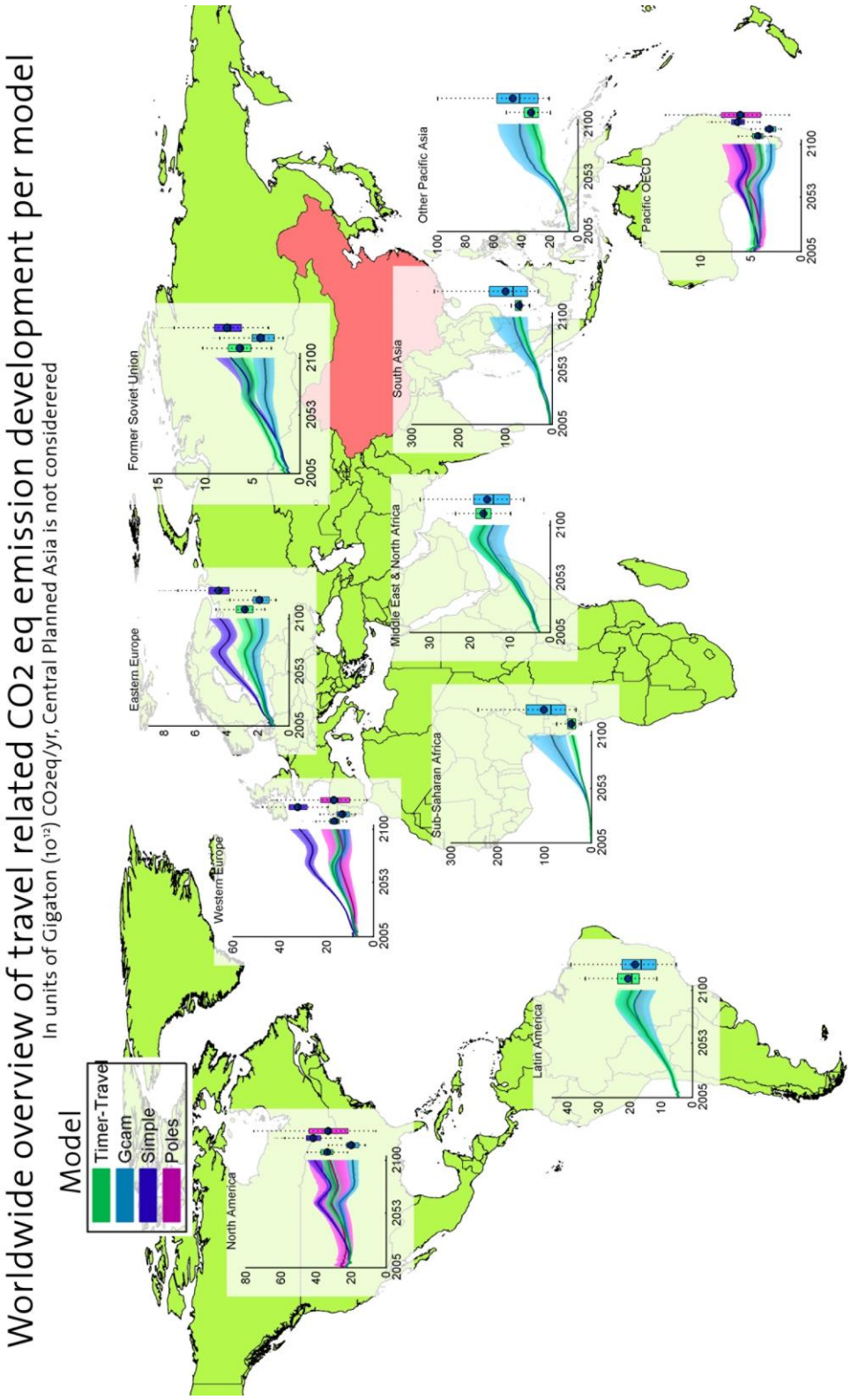


Figure 127: Probabilistic projections for CO_2 emissions per region

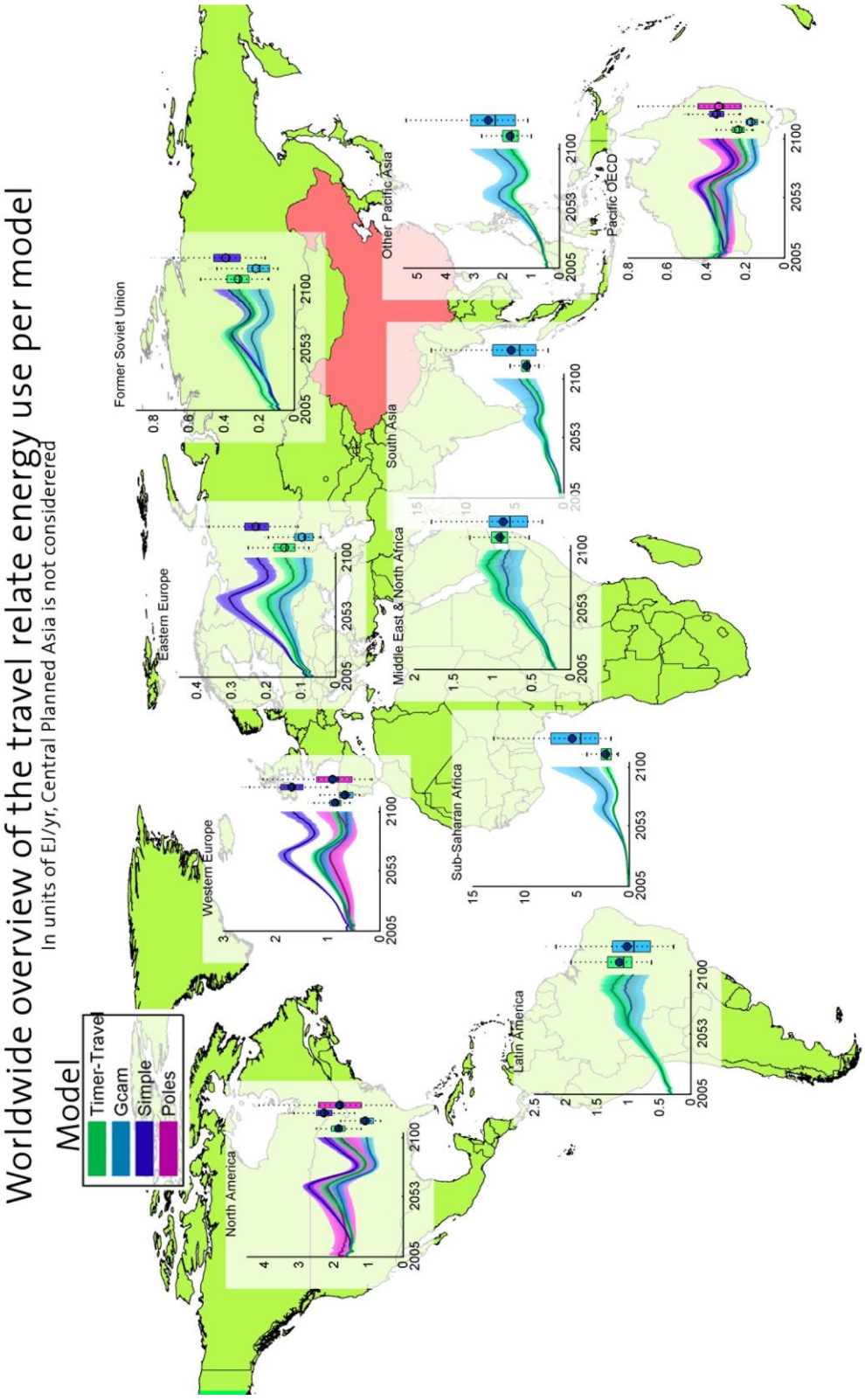

Worldwide overview of the travel relate energy use per model
In units of EJ/yr, Central Planned Asia is not considerered
Model
Timer-Travel
Gcam
Simple
Poles
North America
Western Europe
Former Soviet Union
Eastern Europe
Other Pacific Asia
South Asia
Pacific OECD
Middle East & North Africa
Sub-Saharan Africa
Latin America

Figure 128: Probabilistic projections for energy use per region

A.3 LEFT OUT RESULTS GCAM

A.3.1 REMAINING MODES GCAM

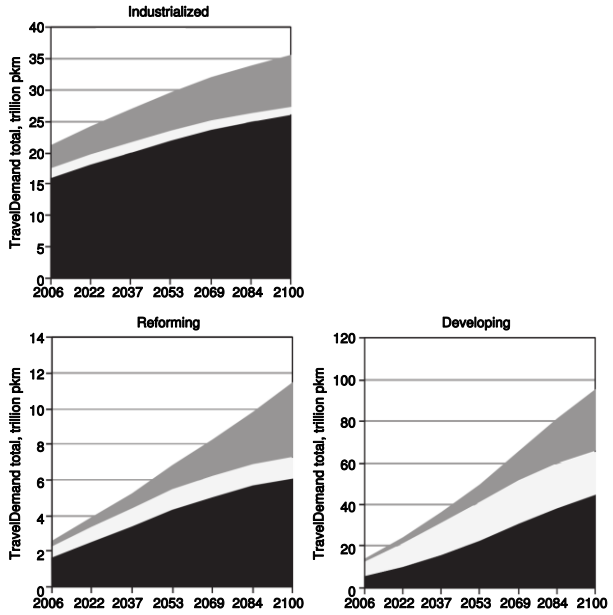


Figure 129: Published results for GCAM

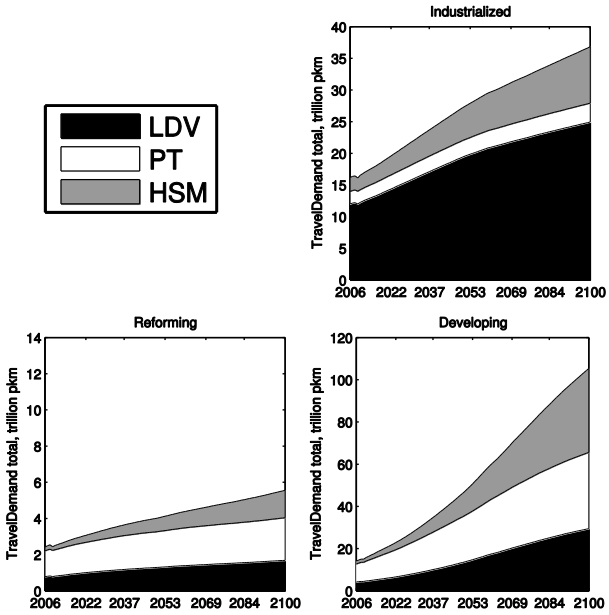


Figure 130: Replicated results for GCAM

A.3.1.1 GCAM WITHOUT UN-MOTIRIZED (ORIGINAL)

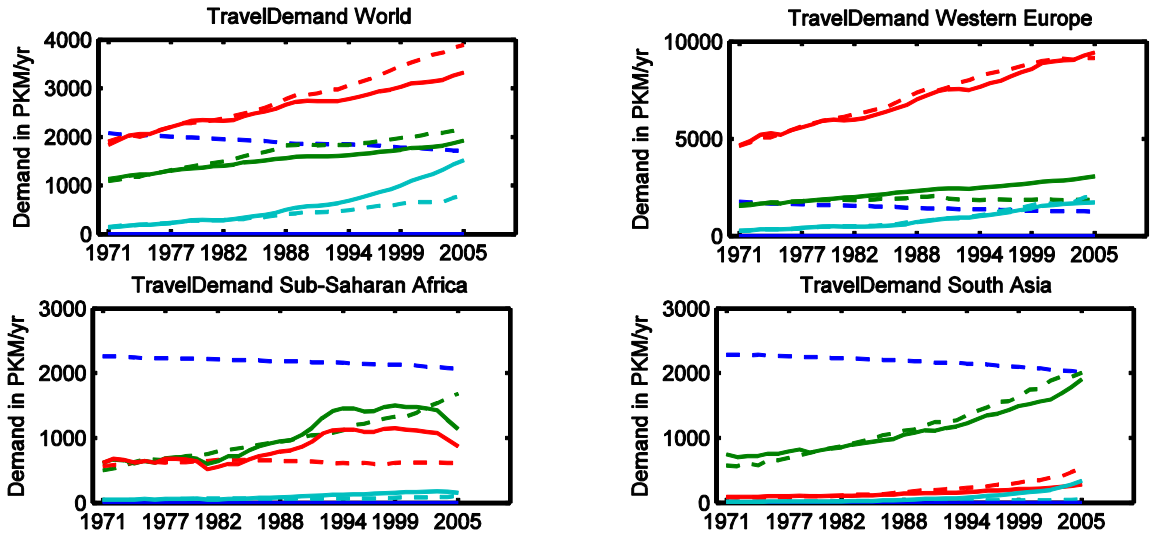
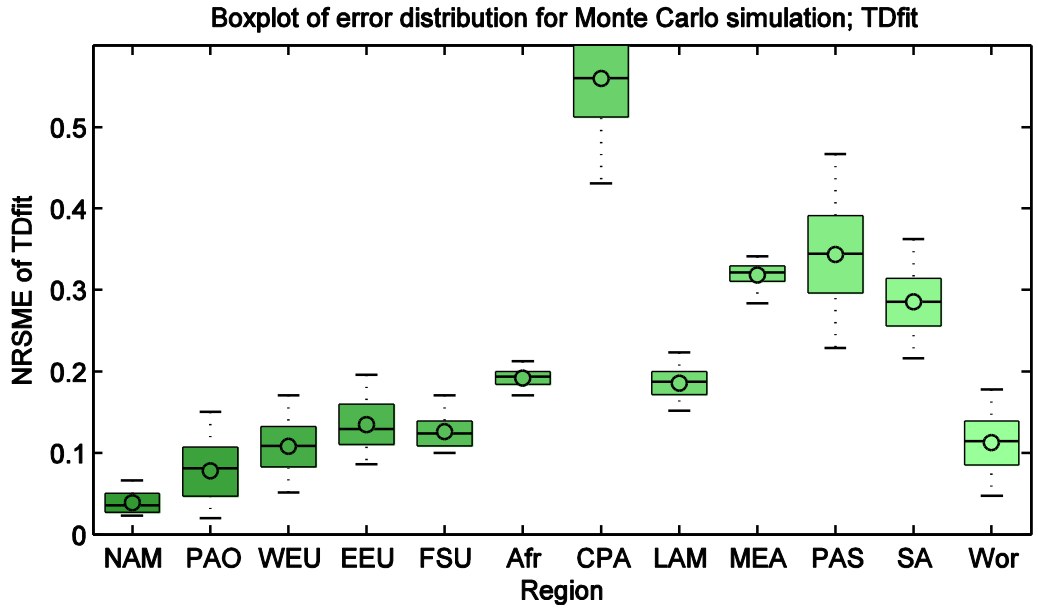
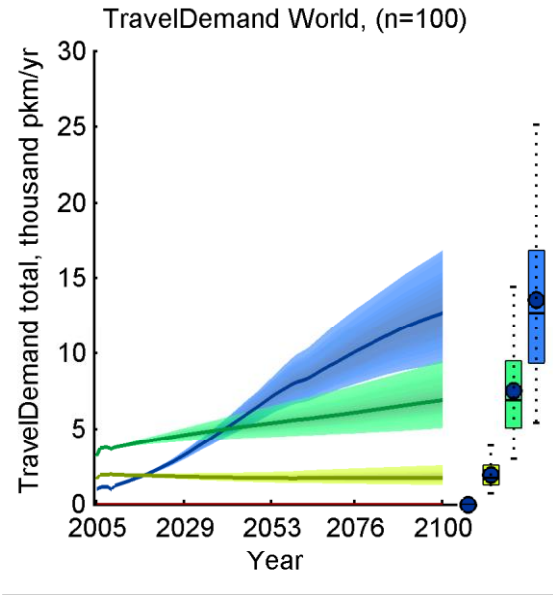
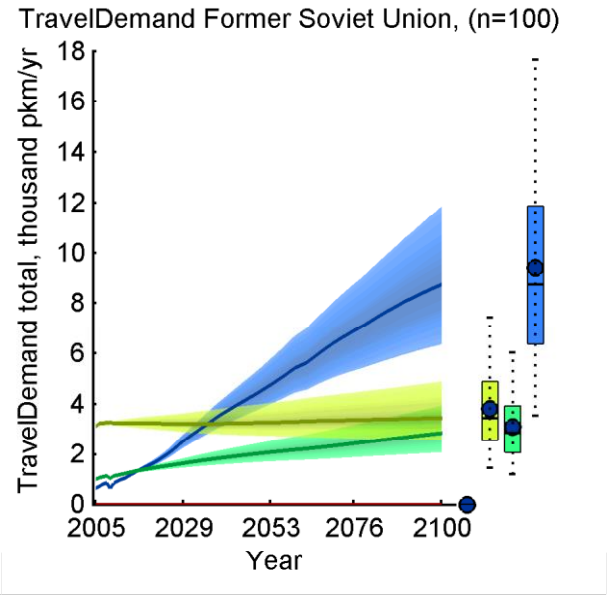
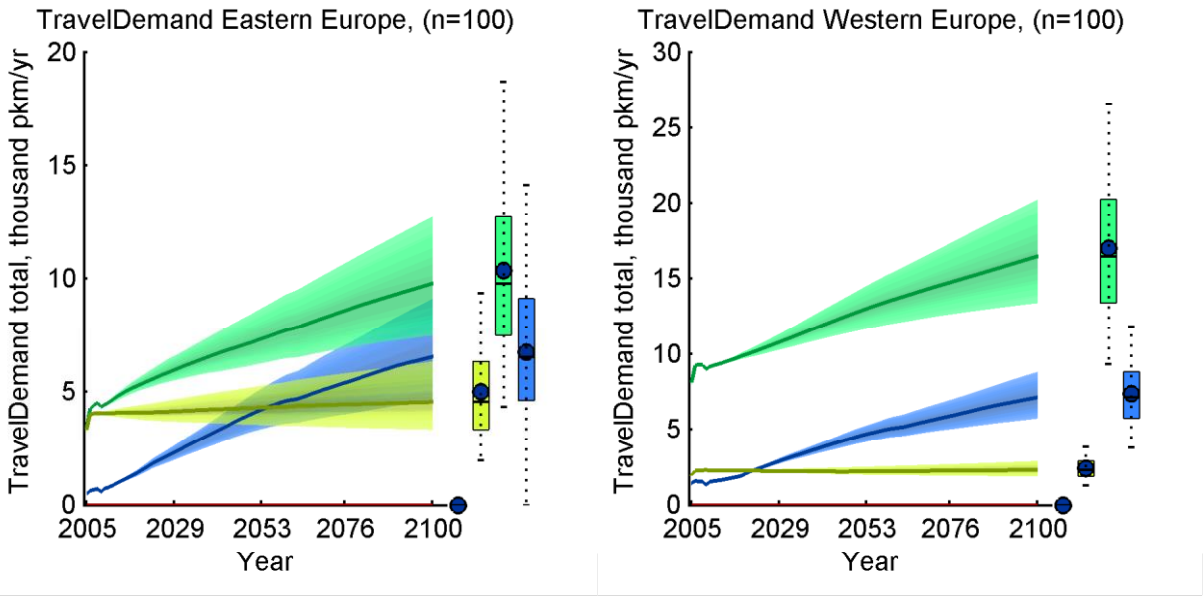


Figure 131: Historic fit of the GCAM model based on calibration over 1971-1980. The respective errors of the fit are 3.7% ,3.4%, 25%, 11%.


Boxplot of error distribution for Monte Carlo simulation; TDfit
NRSME of TDfit
0
0.1
0.2
0.3
0.4
0.5
NAM
PAO
WEU
EEU
FSU
Afr
CPA
LAM
MEA
PAS
SA
Wor
Region


TravelDemand World, (n=100)
TravelDemand total, thousand pkm/yr
0
5
10
15
20
25
30
2005
2029
2053
2076
2100
Year


TravelDemand Former Soviet Union, (n=100)
TravelDemand total, thousand pkm/yr
0
2
4
6
8
10
12
14
16
18
2005
2029
2053
2076
2100
Year



A.4 OTHER MODELS

A.1.1 SCHAFER

This model was dropped after the method was not compatible with the others. The approach is single-run econometrics, which are hard to justify with a probabilistic approach.

The model introduced by henceforth called the Schafer model, consists of combining a large number of regressions and what Schafer calls the target point (TaP). Schafer's model is different from TIMER-travel in that it depends on regressing all of history a once and extrapolating on different formulas for each travel mode. The model predicts for the year 2050 and then works backward to interpolate the pathway taken to get to this point. The lower case letters (*a,b,c,d,e,f,g,h*) represent regression coefficients used in the formulas. A graphic representation of the model can be seen in Figure 132.

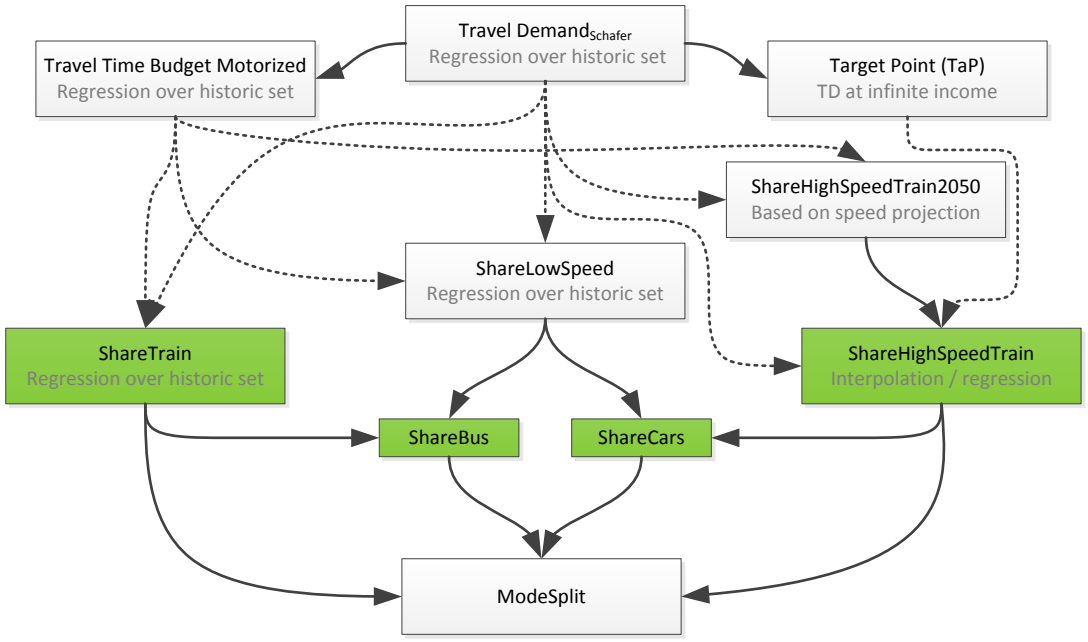


Figure 132: The dependencies of the Schafer model, the green boxes represent each of the four shares

First the total TD per region is determined by:

$$TD_{Schafer,r,t} = \frac{log\left(\frac{I}{g} - h\right)(I)TaP^{1-e}}{log\left(\frac{TaP}{g} - h\right)}$$

With log the natural logarithm, I the income (GDP/capita) and e, g and h regression coefficients. The target point is calculated in the paper as the maximum number of hours traveled per year if the entire TTB were spent on traveling in the fastest modes (airplanes). At a speed of 600 km/hr with 1.1 hr/day, at an income of 240000 $/yr the target point is

$$TaP = Target\ Point \approx 240000 \text{ km/cap/yr}$$

Having calculated the total TD the shares per mode of transportation can be calculated. The TTB, which is assumed to be 1.1 hr/day in the Schafer publication, will first have to be determined (it was assumed 1.2 in). Next TTB is split up into motorized and non-motorized based on another regression over the TD:

$$TTB_{mot,r,t} = a + \frac{b}{(TD - c)^d} = TTB - TTB_{nonmot,r,t}$$

With $a = -\frac{b}{(-c)^d}$ and $b = \frac{TTB_{mot}}{\left(\frac{1}{(TaP-c)^d}\right) - \frac{1}{(-c)^d}}$.

The TTB calculated are next split up into the Share Rail (SR), the Share Bus (SB), the Share High Speed Train and air travel (SHST) and the Share Car (SC). The first two categories belong to the Low Speed modes (SLS). The walking and biking shares (present in TIMER-travel) are not present in this model explicitly, but are assumed to be contained on the TTB_{nonmot}.

$$MS_{total} = MS_{mot} + MS_{nonmot}$$

$$MS_{mot,r,t} = \frac{TTB_{mot,r,t}}{TTB}\{SR, SB, SHST, SC\}\begin{cases} SR = Rail \\ SB = Bus \\ SHST = Air + High\ speed\ train \\ SC = Share\ Car \end{cases}$$

First the share of low speed modes is defined as:

$$SLS = Share\ Low\ Speed = \ Share\ Bus + Share\ Rail = SB + SR$$

Hence it is sufficient to know the share low speed and the share rail to derive the share bus. First the share of rail is calculated as follows:

$$SR_{r,t,i,j,k} = i\left(\frac{1}{\left(TD_{r,t,rail} - j\right)^k} - \frac{1}{(TaP - j)^k}\right)$$

Adding curves at a later stage

With i and k as regression parameters. Next the calculation of the SLS is more complicated, for the authors decide to split it up into two parts. They observe two different characteristics, first for the historic data and then for the projected data. The former is observed to have the characteristics of a 4D polynomial. The latter is regressed such that is fits the last point of the first segment. Hence for the first segment:

$$SLS_{r,t\leq t0,n,o,p,q,r} = n + o\ TD + pTD^2 + q\ TD^3 + r\ TV^4$$

With n, o, p, q, r as regression coefficients; calculated up until the first year of the second segment. The second segment is next defined as:

$$SLS_{r,t>t0,l,m} = \ l\left(\frac{1}{TD - m} - \frac{1}{TaP - m}\right)$$

With l and m regression coefficients. One can next combine these two into the following equation:

$$SLS_{r,t,p,q,r} = \frac{l}{TD_0 - m}\left(1 + \frac{TD - TD_0}{m - TD_0}\right) + p\ (TD - TD_0)^2 + q(TD^3 - 3TD_0^2TD + 2TD_0^3) + r(TD^4 - 4TD_0^3TD + 3TD_0^4$$

With TD0 the TD at the point of convergence. Next the share of high speed train is calculated by using the following equation:

$$SHST_{t=2050,r} = Share\ HighSpeed\ Train\ 2050 = \frac{1 - SB\left(1 - \frac{V_C}{V_B}\right) - SR\left(1 - \frac{V_C}{V_R}\right) - \frac{V_C TTB_{mot} 365}{TD}}{1 - \frac{V_C}{V_{HST}}}$$

With V_c, V_{HST}, V_R and V_B referring to the future mean speed of the relative modes. This determines the share in 2050, from this year the authors project backward (interpolate) to the present day using the following regression:

$$\mathrm{SHST}_{r,t} = \frac{\mathrm{SHST}_{r,t=2050} - 1}{exp\left(\mathrm{e}^{-t(\mathrm{SHST}_{t=2050}-u)}\right) - exp\left(\mathrm{e}^{-t(\mathrm{TaP}-u)}\right)} exp\left(\mathrm{e}^{-t(\mathrm{TD}-u)}\right) + 1 - \mathrm{m}\, exp\left(\mathrm{e}^{-t(\mathrm{TaP}-u)}\right)$$

With *t* and *u* regression coefficients. Lastly the share for cars is derived from the share remaining.

$$\mathrm{SC} = \text{Share car} = 1 - \mathrm{SLS} - \mathrm{SHST}$$

B. APPENDIX: DETAILED MATERIAL

B.1 LISTING OF DATA SOURCES

Listed below in Table 18, Table 19 and Table 20 are the sources for the extensive elasticity investigation made in the present investigation. The sources presented are only the review articles summarizing the values, the sources mentioned there have been consulted for validation purposes but are not presented here.

Vehicle type	Fuel Short	Fuel Long	(Review) Articles
PT on LDV cost		0,22 ± 0,11	(Graham & Glaister, 2004; Litman, 2012a)
LDV	-0.15 ± 0.07	-0.41 ± 0.2	(Ajanovic, Dahl, & Schipper, 2012; Chamon, M. Mauro, P. Okawa, 2008; Graham & Glaister, 2004; Hedenus, Karlsson, Azar, & Sprei, 2010; Litman, 2011, 2012b, 2012c; MIT Energy Initiative, 2008; Sparke, n.d.; Storchmann, 2005)
PT	-0.1 ± 0.05	-0.3 ± 0.1	(Currie & Phung, 2006; Litman, 2011, 2012b; Paulley et al., 2006)
Air	-0.15 ± 0.08	-0.3 ± 0.15	(Airbus, 2012; Federal Aviation Authority, 2012; IATA Consulting, 2011; IETA, 2007; Wohlgemuth, 1998)
Vehicle total	-0.21 ± 0.06	-0.56 ± 0.09	(Litman, 2012a; Rodney, Falvey, & Gemmell, 1996; Schafer, Heywood, Jacoby, 2010)
Average	-0,16 ± 0,07	-0,15 ± 0,13	
Individual sources	20	24	

Table 18: Fuel price elasticities

	Income Short	Income Long	(Review) Articles
LDV ER		1+/-0.25	(Chamon, M. Mauro, P. Okawa, 2008; Paulley et al., 2006)
LDV	0.54 ± 0.29	1.01 ± 0.3	(Ajanovic et al., 2012; Chamon, M. Mauro, P. Okawa, 2008; Graham & Glaister, 2004; Hedenus et al., 2010; Litman, 2011, 2012b, 2012c; Sparke, n.d.)
PT	-0.35 ± 0.05	-0.59 ± 0.26	(Currie & Phung, 2006; Litman, 2011, 2012b; Paulley et al., 2006)
Air		1.86 ± 0.6	(Airbus, 2012; Federal Aviation Authority, 2012; IATA Consulting, 2011; IETA, 2007; Wohlgemuth, 1998)
Vehicle total		0.98 ± 0.1	(Litman, 2012a; Schafer, Heywood, Jacoby, 2010)
Average	0.24 ± 0.17	0.76 ± 0.29	
Individual sources	3	15	

Table 19: Income elasticities

	Price Short	Price Long	(Review) Articles
LDV	-0.29	-0.81	(Ajanovic et al., 2012; Chamon, M. Mauro, P. Okawa, 2008; Graham & Glaister, 2004; Hedenus et al., 2010; Litman, 2011, 2012b, 2012c; Sparke, n.d.)
PT	-0.35 ± 0.15	-0.79 ± 0.13	(Currie & Phung, 2006; Litman, 2011, 2012b; Paulley et al., 2006)
Air		-0.74	(Airbus, 2012; Federal Aviation Authority, 2012; IATA Consulting, 2011; IETA, 2007; Wohlgemuth, 1998)
Vehicle total		-0.75 ± 0.25	(Litman, 2012a; Schafer, Heywood, Jacoby, 2010)
Average	-0,34 ± 0,15	-0,77 ± 0,15	
Individual sources	6	20	

Table 20: Travel Price Elasticities

B.2 MATLAB MODELING FRAMEWORK

The Matlab framework created is open source under the MIT open source license and available from https://globaltravelmodels.codeplex.com/. All code can hence be used freely under mention of the original authors. The code contains parallel computing, dozens of plots, a file handling system and is fully operable from a user interface, with popup menus. It can also export professional graphs, and handle multiple output files at once. The description here is short, and if the reader considers using the framework the model should simply be run and comments in the code be read.

From a user point of view the program starts with the Startmodel command which opens a user interface which can be used to start simulations. The gui contains a selection of selection tools, most of them are commented in the code, or can simply be tested by running a model. If models have been run they are saved to a location of choosing, if the default folder is used the load last run button can load the data. The saved mat files can then be loaded by post processing. See figure below.

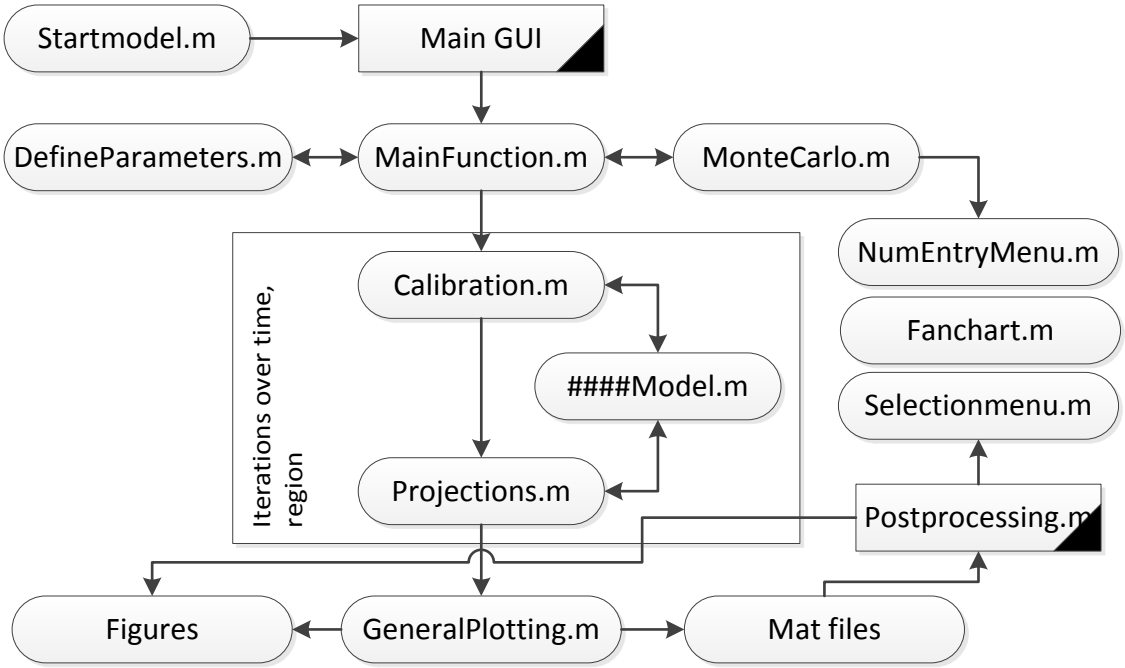


Figure 133: Overview of modeling framework

B.3 FULL FORMULATION POLES

The POLES model is based on fuel price elasticities δ, TD to income elasticities ϵ and equipment rate to income elasticities ERδ, all defined per region and generally per mode. The latter elasticity is modified based on the distance to a saturation level reducing it to zero near saturation levels, ERSAT (vehicles/capita).

$$ER\delta_{r,t,car} = ER\delta 0_{r,car} * \left(\frac{ER_{r,t,car}}{ER_{r,t-1,car} - ERSAT_{r,car}}\right)^{0.05} [-] \quad (20)$$

Next the equipment rates for cars, and motorcycles, mot, are calculated, their values in the first year of the simulation are based on the dataset described in section 2.4.3. The motor cycles are omitted in the model comparison due to lacking data in the Schafer set (Schafer, Heywood, Jacoby, 2010), but taken into account in the discussion section. For cars the equipment rate depends only on the ϵ, notice that this equipment rate is never allowed to shrink relative to the last year, even if there is economic recession.

$$ER_{r,t,car} = ER_{r,t-1,car}\left(\frac{I_{r,t}}{I_{r,t-1}}\right)^{ER\epsilon_{r,t,car}} \; with \left(\frac{I_{r,t}}{I_{r,t-1}}\right) \geq 1, else \; 1 \left[\frac{vehicles}{1000 \; cap}\right] \quad (21)$$

For motor cycles the approach is different, it behaves by first increasing and then decreasing. Two variables, similar to elasticities, determine the rate of increase and decrease, α and β.

$$ER_{r,t,mot} = ER_{r,t-1,mot}\left(\frac{1}{I_{r,t}} - \frac{1}{I_{r,t-1}}\right)^{\alpha} + \left(I_{r,t} - I_{r,t-1}\right)^{\beta}\left[\frac{vehicles}{1000\ cap}\right] \quad (22)$$

For cars the average vehicular travel demand VTD (vkm/yr) are used to determine how many vehicle kilometers are traveled by an average car. This is based on fuel prices FP (\$/pkm), aforementioned elasticities and a time trend TT. The price elasticities in this formulation are dependent on the present year and two years back, last year (t-1) is weighed $2/3^{rd}$, the year before that (t-2) $1/3^{rd}$. δ is negative, ϵ is positive.

$$VTD_{r,t,car} = VTD_{r,t-1,car}\left(\frac{FP_{r,t,car}}{FP_{r,t-1,car}}\right)^{\frac{2}{3}\delta_{r,car}}\left(\frac{FP_{r,t-1,\mathrm{car}}}{FP_{r,t-2,\mathrm{car}}}\right)^{\frac{1}{3}\delta_{r,car}} \left(\frac{ER_{r,t,car}}{ER_{r,t-1,car}}\right)^{ER\delta_{r,t,car}} * \left(1 + TT_{r,car}\right)\left[\frac{vkm}{yr}\right] \quad (23)$$

These formulae give TD for LDV when combined with a parameter for motorcycles EQ. Representing the number of TD per motorcycle compared to cars, which is given by the authors as 0.2, i.e. motorcycles travel 5 times less than cars.

$$TD_{r,t,LDV} = VTD_{r,t,car}\left(ER_{r,t,car} + EQ_{r,t,mot}\ ER_{r,t,mot}\right)\left[\frac{pkm}{cap\ yr}\right] \quad (24)$$

The formulation for public transportation is next calculated. It is directly dependent on the TD the year before, but otherwise similar to the calculation for cars. In the original model this mode further disaggregated into rail and bus, but this is not required for the present investigation. δ and ϵ are both positive, since they depend on the fuel price for cars.

The way these fuel price elasticities are formulated hence behave similarly to cross elasticities. Car equipment rate depends negatively on a high fuel price, whereas public transport TD will increase with a high price. Hence PT replaces cars with increasing fuel prices and cars replace PT with decreasing fuel prices.

$$TD_{r,t,PT} = TD_{r,t-1,PT}\left(\frac{FP_{r,t,car}}{FP_{r,t-1,car}}\right)^{\frac{2}{3}\delta_{r,PT}}\left(\frac{FP_{r,t-1,car}}{FP_{r,t-2,car}}\right)^{\frac{1}{3}\delta_{r,PT}} \left(\frac{I_{r,t}}{I_{r,t-1}}\right)^{\epsilon_{r,PT}}\left(1 + TT_{r,PT}\right)\left[\frac{pkm}{cap\ yr}\right] \quad (25)$$

Lastly the TD for air is defined. There is a saturation on σ that is similar to the saturation on the car equipment rate elasticity, except the saturation TDSAT (pkm/cap/yr) is on the TD for air. For air σ is negative and ϵ positive, with historically declining prices this has historically meant an increase in the TD for air.

$$\epsilon_{r,air,t} = \epsilon 0_{r,air} * \left(\frac{TD_{r,\mathrm{t},air}}{TD_{r,t-1,air} - TDSAT_{r,air}} \right)^{0.05} [-] \quad (26)$$

$$TD_{r,t,HSM} = TD_{r,t-1,HSM} \left(\frac{FP_{r,t,air}}{FP_{r,t-1,air}} \right)^{\delta_{r,air}} \left(\frac{I_{r,t}}{I_{r,t-1}} \right)^{\epsilon_{r,t,Air}} \left(1 + TT_{r,air}\right) [\frac{pkm}{cap\ yr}] \quad (27)$$

Values for the POLES model were acquired through literature research and can be seen in Table 8 and a full description can be found in the appendix. Values from the author were not available. Literature research on the transportation elasticities frequently limits itself to regional studies in OECD regions. It is certainly not possible to separate the variables into individual regions, the databases that exists with this information are incomplete (Dahl, 2011). Information exists on the difference between developing and industrialized nations (IETA, 2007) and the development of elasticities using econometrical methods independent of country (Chamon, M. Mauro, P. Okawa, 2008; Dargay et al., 2007) but not enough information exists to do this for every variable, for every region. The values are therefore based on investigations in the UK, US and Europe.