ETH
Eidgenössische Technische Hochschule Zürich
Swiss Federal Institute of Technology Zurich

Spatio-Temporal Estimation of Global Plastic Waste Inputs to the Ocean from 1990 to 2015

Master Thesis

K. Lang

24.09.2018

Advisors: Dr. Charlotte Laufkötter, Dr. Meike Vogt
Department of Environmental Sciences, ETH Zürich

Spatio-Temporal Estimation of Global Plastic Waste Inputs to the Ocean from 1990 to 2015

K. Lang[1]*

Abstract
Marine plastic pollution is an increasing global concern, as plastic particles are found in the most remote parts of the ocean and pose various threats to marine life. First estimates quantify the contemporary coastal or riverine plastic input at the country-level, but little is known about the spatio-temporal distribution of its sources and sinks at the global scale.
We use machine learning algorithms to estimate the global plastic input to the ocean from coastal sources and riverine input for the years 1990-2015 on a 0.1x0.1° grid. Using neural network and random forest models, we predict the waste generation per country from socio-economic factors such as population density or energy consumption. We then apply population size, waste management and waste composition, as well as distance-based probabilities of land and river transport to derive the plastic flux to the ocean. Additionally, the effect of several waste management practise scenarios on plastic fluxes as well as a future projections assuming a business as usual scenario for 2030 and 2050 are assessed. Finally, we determine the fate of the floating marine plastic debris using a Lagrangian particle simulator driven by geostrophic current data.
In line with previous findings, predicted global plastic input increased from 13.4 Mt in 1990 to 18.7 Mt in 2015 with 94% and 6% entering via the coasts and rivers respectively in 2015. While middle-income countries (representing 58% of the world's population) account for 90% of the global marine plastic input in 2015, high-income countries (31% of the population) account for 9% and low-income (11% of the population) for 1%. In 2015, most plastic enters the oceans in Asia (84%) while Europe and Africa account for 5% each, North and South America for 3% each and Oceania for 0.2%. Our models also suggest an increase of marine plastic input in several African, East Asian and South American countries in the time period observed while European and Northern American countries generally show a decrease. Our results show the temporal and spatial development of marine plastic debris generation and the influence of the socio-economic factors. This helps to assess where possible measures to reduce plastic pollution will be most effective.

[1] *Department of Environmental Sciences, ETH Zürich, Switzerland*
***Corresponding author**: langk@student.ethz.ch

Contents

1. Introduction

The worldwide production of plastic goods rose exponentially since its introduction to the public after World War II, from 2 Mt in 1950 to 380 Mt in 2015 (average yearly growth rate of 8.4%) [1]. Properties such as light weight, durability, bio-inertness, oxygen/moisture barrier and low costs make plastic an ideal packaging material [2]. Thus, 44.8% of the worldwide plastic resin production is used for packaging materials which have a mean lifetime of half a year [1]. Almost half of the 302 Mt of plastic waste produced in 2015 [1] was therefore single-use plastic. With increasing plastic waste production, an increasing contamination of plastic waste in the natural environment can be expected [3].

Over the past five decades, the plastic waste input to the oceans has continuously risen [2, 4]. While the first report of plastic in the oceans dates back to 1972 [5], estimates for 2010 suggest that 4.8-12.7 Mt entered the ocean from the coast [4]. As of today, plastic has been found in the most remote areas of the ocean and in various sizes [6], ranging from the nano ($<$100 nm) to the micro ($<$5 mm) and meso ($>$5 mm) scale [2, 6]. This global distribution of marine plastic debris has severe effects on marine biota.

It is estimated that over 90% of the damage caused to marine animals by human waste is due to plastic pollution [7]. Mesoplastics, including lost fishing gear, six-pack drink holders or balloons, can entangle animals. This makes them more vulnerable to predators or can lead to starvation, strangulation or drowning due to exhaustion [8]. Animals such as sharks, sea turtles, sea mammals and birds are especially vulnerable to entanglement [8]. Furthermore, mesoplastics are often ingested by various types of animals. Kühn et al. [8] documented ingestion of plastics in 44% of 446 observed birds, turtles and sea mammals and Auman et al. [9] found plastic in 97.6% of stomachs of chicks of the Laysan albatross. Plastic ingestion can block or damage the gastrointestinal track, which may lead to death [8]. Additionally, plastic pieces can block parts of the stomach, which can lead to reduced digestive processes and a feeling of satiation, which results in malnourishment, dehydration and, eventually, death [8]. In addition to the threats of mesoplastics, micro- and nanoplastics can have deleterious effects on many organisms. Due to their small size they can be ingested, as observed in zooplankton such as *calanus helgolandis* [10] or *daphnia magna* [11, 12], algae such as *scenedesmus obliquus* [12], invertebrates such as the fiddler crab (*uca rapax*) [13], sea urchins such as *paracentrotus lividus* [14], sea cucumbers (*holothuroidea*) [15] or corals such as *acropora*, *pocillipora* and *porites* [16]. Besides these direct physiological and physical effects on biota, plastic also poses a chemical risk.

Marine plastic debris consists of a complex mixture of chemicals, including by-products such as polycyclic aromatic hydrocarbons (PAHs) or absorbed compounds such as polychlorinated biphenyls (PCBs), polybrominated diphenyl ethers (PBDEs) and heavy metals such as Ni and Pb [17]. 78% of the priority pollutants (a list containing 126 well known toxic pollutants in water which is often used for regulations) as listed by the United States Environmental Protection Agency were associated with marine plastic debris [18]. Bioavailability upon ingestion has been shown in several marine animals such as whales, sharks, seabirds and fish [17]. The ingestion of plastic particles can furthermore lead to the accumulation of deleterious pollutants in the animal itself and the food chain [17]. This can cause various sub-lethal effects in smaller animals or poisoning of marine mega fauna and apex predators such as sharks [8, 17–21]. Additionally, plastic particles may also create an ecological problem by acting as vectors for invasive species [22]. The global economic cost due to damage in marine ecosystems caused by plastic pollution has been estimated to amount to 13 billions USD per year [7]. Because of the mentioned risk and threats of marine plastic debris, the increasing marine pollution has risen awareness. In the scientific community and among the politicians and the public, efforts are made to understand and solve the problem of plastic pollution.

Political efforts to reduce plastic pollution include plans for product bans and for a 90% collection rate of single-use drink bottles in the EU [23] or the recent total ban of plastic bags in Kenya [24]. Meanwhile, the scientific community tries to gain further insight on the governing processes of plastic pollution. The number of publications on marine plastic debris were exponentially increasing in the last decade (compare Fig. 1). However, the sources, sinks, amount and distribution of plastic particles in the ocean are still barely constrained. Open research questions include the limited understanding regarding the pathways of terrestrial plastic to the marine environment [25], the effects of wind and currents on the floating behaviour of litter items [26], the degradation processes of plastics in the ocean [25, 26], or the missing understanding of (future) trends of plastic debris [27]. Nonetheless, several studies have created insights by providing estimates of plastic input to the ocean and current amount of marine plastic debris (compare table 1). However, there is a large difference in the modelled estimates and measurements.

Modelling approaches to predict the amount of marine plastic debris based on the total plastic production [1] or on the municipal solid waste (MSW) production [4, 33, 34] estimate yearly plastic inputs on the order of millions of tons. On the other hand, estimates based on ocean surveys predict the total amount of plastic floating at the surface only on the order of thousands of tons. This difference between modelled estimates and measured concentrations is known as the *missing plastic problem* [35] and has been discussed since 2004 [3]. Furthermore, several ocean regions (including the Eastern Pacific and the North Sea) do not show an increase in surface plastic concentrations since the 1980's [25, 36]. This is surprising because the production of plastic waste has increased drastically since and we thus could expect an increased flux to the ocean and thus higher oceanic plastic concentrations [1, 3, 37]. The *missing plastic problem* and the lacking increase in the concentrations thus imply that the ocean surface is not

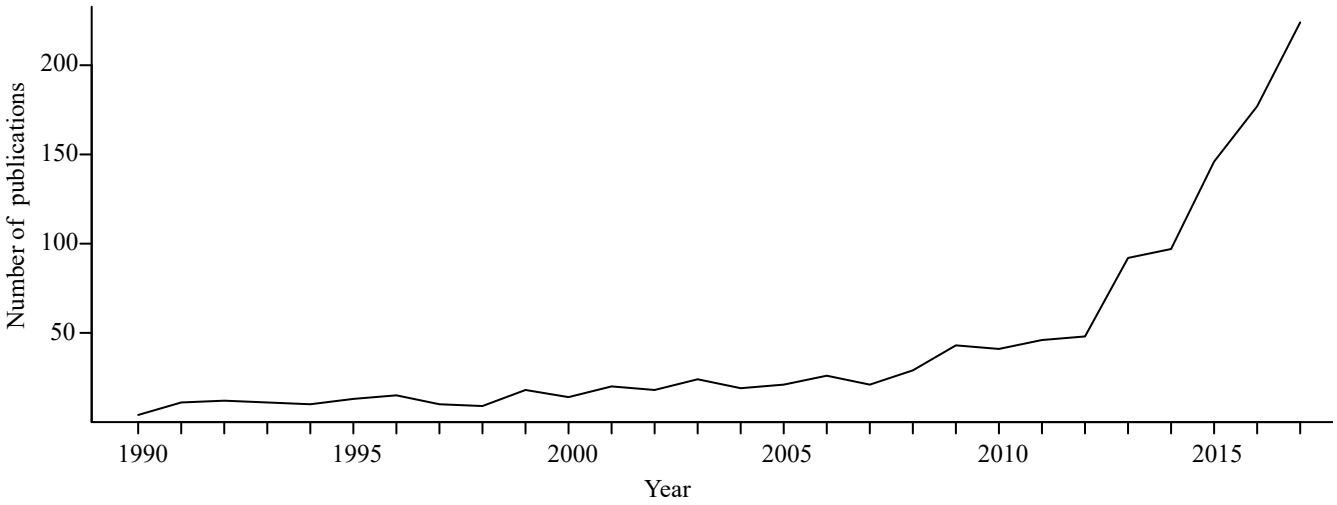

Figure 1. Number of publications listed on Thomson Reuters Web of Science when searching for the terms "ocean" and "plastic" in both title and topics (search query: *TI=(ocean AND plastic*) OR TS=(ocean AND plastic*)*). As of 30. May, 122 publications were listed in 2018 that are not included in the graph. In total, 2'001 publications have been listed from 1968 until 30. May 2018.

Table 1. Summary of estimates on the amount of plastic debris in the worlds oceans. We present predictions based on measurements and predictions based on model approaches. It is obvious that the modelling approaches predict higher plastic masses in the ocean that what the measurements suggest. This difference is known as the *missing plastic problem*.

	Location	Estimate of plastic in the ocean		Description	Source
		Particles	Mass		
Measurements	All oceans	$5.25 \cdot 10^{12}$ $4.85 \cdot 10^{12}$ (microplastic)	268.9 kt 35.5 kt (microplastic)	Estimate based on 24 expeditions in 2007-2013	[28]
	All oceans		6.6-35.2 kt	Estimates based on 2010 Malaspina circumnavigation measurements	[25]
	All oceans	$15\text{-}51 \cdot 10^{12}$ (microplastic)	93-236 kt (microplastic)	Estimates on microplastic particles based on a statistical model incorporating >11'000 observation (including [28] and [25])	[29]
	Southern Ocean	100'000 particles/km^2		Measurements from 5 net tows on an expedition in 2016	[30]
	North Pacific Central Gyre	334'271 particles/km^2		Measurements from 11 net tows in August 1999	[31]
Model estimates	Coastal regions		4.8-12.7 Mt/year	Model including mismanaged MSW of coastal populations in 192 coastal countries	[4]
	All oceans	$1.1\text{-}3.6 \cdot 10^{18}$	45-129 kt	Model based on multiple boat and aircraft surveys	[32]
	All oceans		1.15-2.41 Mt/year	Modelled yearly riverine inputs based on mismanaged MSW and river flow model	[33]
	All oceans		0.41-4 Mt/year	Modelled yearly riverine inputs based on mismanaged MSW data from [4] extended with river catchment sizes	[34]

the ultimate destination of plastic that enters the ocean and that unknown sinks exist, that rapidly remove plastic debris from the ocean surface [25, 28, 36].
Possible removal process of plastic from the ocean surface include degradation, mineralization and sedimentation [2, 25]. As already mentioned above is our understanding of the removal processes of marine plastic debris still limited. However, several mechanisms that could remove plastic from the surface have been suggested. Here we present an overview of these removal processes and their mode of action:

- **UV degradation:** The UV-B radiation in sunlight initiates a photo-oxidative degradation in many common plastics [2]. UV degradation is orders of magnitude faster compared to other types of degradation such as hydrolysis [2]. The process adds oxygen rich functional groups to the plastic and lowers its molecular weight [2] which makes the plastic less stable. Degraded plastics are more prone to be fragmented by physical forces such as waves. However, the light-induced oxidative degradation is severely less effective on plastics floating in water due to lower temperatures and a decreased oxygen availability compared to plastics out of the water [2]. Furthermore, UV radiation is rapidly absorbed by the water column and the process is thus only relevant at the water surface [2]. Thus, UV degradation can decrease the size of plastic particles.
- **Beaching:** As UV degradation is hindered in water and other degradation processes such as hydrolysis are much slower, UV degradation on beaches likely plays an important role in breaking down mesoplastics [2]. Due to the low specific heat of sand, plastic on beach surfaces can reach temperatures of around 40°C [2]. As the rate of the oxidative degradation doubles when rising the temperature by 10°C (Arrhenius equation), plastics on the beach will degrade even more rapidly [2]. Additionally, the mechanical damage of waves can break down mesoplastics [38]. Thus, beaching can also decrease the size of plastic particles.
- **Biofouling:** Plastic floating in the water will rapidly be colonized by bacteria, diatoms or barnacles. This process is called biofouling and increases the density of the particle [2]. An increased density will result in sinking of the plastic particle. Thus, biofouling removes plastic from the ocean surface. Additionally, a colonization by microbes can also start the process of biodegradation [2]. This describes the breakdown or mineralization by organisms [2]. As the rate of biodegradation is dependant on the molecular weight of the plastic, already UV degraded microplastic particles will be faster biodegraded than mesoplastics [2]. Thus, colonization can also lead to removal of microplastics due to mineralization. However, by blocking the light, biofilms also reduce UV degradation which is the faster degradation process on the ocean surface than biodegradation [2]. As the rate of biodegradation is dependant on the molecular weight of the plastic, already UV degraded microplastic particles will be faster biodegraded than mesoplastics [2]. However, UV degradation will only decrease the molecular weight down to 10^3-10^4 g/mol whereas microbial biodegradation requires molecular weight of 500 g/mol or lower to be efficient [2]. Thus, many plastics are relatively fast degraded to the microplastic range but will not be mineralized at similar rates. Estimated mineralization rates vary from 10-20 years for plastic bags, 50 years for styrofoam cups and up to 600 years for fishing lines [39]. Thus, biofouling can lead to fast rates of removal due to increased densities as well as to slow rates due to mineralization. On the other hand, biofouling decreases UV degradation.
- **Grazing:** Digestion and excretion of plastic particles can further degrade them. Microplastic particles ingested by the Antarctic krill (*Euphausia superba*) were found to be fragmented into pieces of less than 1 µm, small enough to cross physical barriers [40]. Additionally, plastic fragments can be bound in faeces and sink [41]. Also in the case of death of these organisms, the plastic will sediment together with the biological material [25, 28, 41].

These processes could help to explain the sinks of marine plastic debris from the ocean surface. However, due to the several mechanisms and their influence on each other, estimates on the longevity of floating plastic are still unreliable and include high uncertainties [26]. Estimates of the degradation rates of marine plastic debris (based on drift trajectories of plastic particles, stage of the rafting community, size of rafting organisms and the degradation stage) range from several weeks to centuries, depending on the type of plastic and the environmental conditions [26]. However, by combining measured plastic concentration data and modelled plastic input with quantitative process descriptions, Koelmans et al. [42] estimate that floating plastic gets removed from the ocean surface layer in about three years due to the processes mentioned above. This implies that 99.8% of all plastic that entered the ocean from 1950-2016 has settled [42] which could help to explain the *missing plastic problem*.
Additionally to the unknown sinks, the methodology used to quantify the amount of floating plastic debris in the ocean might accentuate the *missing plastic problem*. Andrady [2] highlights that the common quantification of floating plastic debris using nets neglects plastic particles in sediment and mid-water and thus seriously underestimates the total plastic in the ocean. Brunner et al. [43] suggest a multiplication factor of 3-13 to estimate the total buoyant microplastic marine debris from measurements using nets tows. Additionally, measurements often focus on the known aggregation zones in the subtropical gyres and large parts of the ocean surface outside these zones remain unsurveyed [29].
In order to better understand the sources and sinks of marine plastic debris, there is a need to gather more information on the plastic input (quantity, location and temporal variations),

the possible degradation and mineralization mechanisms and mechanisms that lead to settling of the particles. These insights could help to better explain the *missing plastic problem*. With this thesis, we aim to create a spatio-temporal prediction of plastic fluxes from land based sources to the oceans. We furthermore want to couple this model with an ocean current simulation to predict the dispersal of the marine plastic debris. To reach this goal, we predict the total production of MSW (MSW_T) on country-level for the years 1990-2015. We therefore use machine-learning models (MLM) trained on socio-economic predictor variables (PV) such as greenhouse gas production or the fraction of urban population and test it on known MSW collection data (MSW_C) from 78 countries. We then further process this data by applying information on the waste collection system ($F_{collection}$), the fraction of plastic in MSW_C ($F_{plastic}$) and the MSW treatment to estimate the amount of plastic waste that enters the environment (P_E). By further applying gridded population density data and a waste flow model (WFM) for the inadequately managed plastic waste, we create a geospatial estimate of plastic waste entering the ocean (P_O) on a global scale. This is then coupled with a Lagrangian particle advection simulator to predict the dispersal pathways of the floating plastic debris entering the ocean as well as to predict plastic concentrations in the known aggregation zones. These include the subtropical gyres as well as highly populated and enclosed seas, gulfs, and seas [28].
This approach includes several assumptions for which we will present reasoning in the following paragraphs. We also indicate hypotheses and open research question that we will investigate in this thesis.

By creating a model that is valid for each country, we assume that the selected set of PVs includes all the drivers that explain the MSW generation. Therefore, it is important that this set of PVs includes the drivers that have been found to affect the MSW production. Next, we introduce our selection of six PVs. We indicate for which PVs we found links to MSW production in the literature or we explain, why we expect the PV to influence the MSW production.
GDP per person: MSW generation is known to be monotonically rising with increasing GDP per capita [44]. Additionally, countries with higher GDP also show a lower fraction of MSW deposition in landfills and a higher recycling rate of cardboard and paper [44]. Therefore, GDP does not only include the tendency of industrialized countries to produce more MSW but also the capacities for a more efficient MSW management system compared to non-industrialized countries [44].
Share of urban population: The urbanization of a country was found to have a positive effect on MSW generation rates [45]. People living in cities have a better access to services and thus consume more goods, which increases MSW production [45].
Fraction of population aged 0-14 years: The proportion of children in a population was found to have a negative influence on MSW generation [45]. This could be explained by the fact that children negatively affect the amount of the households income spent on consumer goods.
Fraction of households with access to electricity: We assume that households without electricity access produce less MSW compared to households with access to electricity. While the fraction is usually 100% in high-income countries, Sub-Saharan low-income countries show fractions <50% which indicates either the remoteness and/or poverty of many households. We thus expect a positive link between higher electricity access and MSW production.
Energy consumption per person: We assume that the amount of energy that is consumed is positively linked to MSW generation. We reason that people who consume more energy likely are wealthier and can thus consume more goods than people that use energy more sparsely.
Greenhouse gas production per person: In addition to the energy consumption (which accounts for ∼25% of the global CO_2 production [46]), the greenhouse gas production includes further processes of the society, such as transportation of people and goods, manufacturing or agriculture [46]. Here, we use this PV as an indicator of the economic strength of a country. We assume that a higher greenhouse gas production is positively linked to MSW production.

Figure A.1 in the appendix shows that the described trends between these PVs and MSW_C was also found in our data. The data is further described in the sections 2.1 and 2.2. Additional PVs such as policy factor or public attitudes are reported to have a significant effect on MSW generation [44, 45]. However, these predictors are hard to quantify and data is sparsely available on the temporal scale. Furthermore, the effectiveness of the same policy measure might be different among countries [45]. Thus, these parameters are not included in our PV set.

The selection of the PVs, as described above, led us to our **first hypothesis**:
Modelling the amount of MSW production of each country is expected to be possible using the PVs GDP, share of urban population, fraction of population aged 0-14 years, fraction of households with access to electricity, energy consumption and greenhouse gas emissions.

By predicting the plastic waste production using the MSW generation, we assume that the production of plastic waste is linked to the production of MSW. Next we provide reasoning for this relationship by showing the trends in MSW productions and the MSW disposal practices and their effect on plastic waste production.
As both the worldwide affluence and population have risen continuously in the last decades, the production of MSW has followed this trend [37, 47] and reached a total of 3.5 Mt/day in 2012 [37]. For 2025, estimates of the World Bank predict a total of 6.1 Mt/day [37]. With the total MSW production

rising, a peak in global waste production is expected to occur likely after 2100 at up to 10.9 Mt/day [47]. Nonetheless, several high-income countries could have already reached their peak waste production and the future MSW production is strongly dependant on the economic and population growth in Asia and Africa [47] as well as on other developmental factors such as legislation or education [45]. The increase in these regions might be problematic as the growth of MSW production in countries with high fractions of inadequately disposed waste might lead to higher loads of MSW entering the environment.

MSW is often adequately managed in high-income countries using a efficient waste collection system and treatments using incineration, recycling, composting or deposition in controlled landfills [37]. On the other hand, many low-income and mid-income countries have a less efficient MSW collection system and can not afford to manage their MSW in an environmentally sound way [37]. Thus, a high percentage of waste is not adequately disposed but discarded into uncontrolled landfills or open dump sites [1]. Where MSW collection systems are not available at all or can not be afforded and regulations are absent or not enforced, the waste will be disposed the most cost efficient way, which often results in local burning or deposition in open dump sites [48]. Such inadequately managed MSW has a strong impact on the environment and public health [37, 44]. Detrimental effects on public health include higher congenital anomalies in the population close to dumps and landfill sites or a higher risk factor for cancer due to airborne pollutants [44]. Furthermore, inadequately managed MSW produces more leachates and greenhouse gases than adequately managed MSW [48] and is therefore contributing to climate change. Estimates range from a 10-15% reduction of global greenhouse gases if solid waste management is improved. However, the impact of inadequately managed MSW is highly dependent of its composition, which varies strongly among countries [37].

While MSW in low-income to middle-income countries consists mainly of organic material (64% and 54% respectively), MSW in high-income countries consists mainly of paper (31%) [37]. While the organic and paper fraction of MSW will readily decompose and valuable parts such as metals are often recollected, the low value plastics in MSW will remain [48]. As plastic is one of the most environmentally persistent types of MSW [6] it can thus easily accumulate in the environment if the MSW is inadequately disposed [6, 28]. The fraction of plastic in MSW is between 8% and 12%, independent of the country's income [37]. We can therefore expect that trends in global MSW production and treatment also have a high impact on plastic waste in the environment.

Our **second and third hypotheses** are thus:
Trends in the global production and disposal of MSW are expected to affect the plastic waste that enters the environment.

The income class of each country is expected to influence the amount of plastic waste due to decreased effectiveness of the waste collection system with decreasing income.

Finally, by coupling our prediction of plastic flux to the ocean with a Lagrangian particle advection simulation including sinks of particles, we can create estimates of the distribution of plastic particles. Circulation models have already been used to estimate plastic concentrations [29]. By performing this analysis we can compare if our predicted concentration patterns of plastic are realistic by comparing the simulation results to measurements.

Our **fourth hypothesis** is thus:
Advecting our predicted plastic fluxes to the ocean in an Lagrangian particle advection model will result in similar plastic concentration patterns as in the observed aggregation zones.

In the following chapter we will present the methodology used to reach the set goals and check our hypotheses.

2. Methods

In this chapter we present the methods that we used to predict the flux of plastic to the ocean (P_O). Figure 2 gives a graphical representation of the several steps needed to predict P_O, which will be explained in the following sections. We first present the data that we collected for the predictor variables (PV) that we used to model the amount of collected waste (MSW_C) as well as the data including information on the waste collection system and waste treatment (used to predict the MSW_T) in section 2.1. We further present the correlations and trends among the PVs and with the target variable in section 2.2. This is followed by a description of the modification of said data (to allow use in the machine learning models (MLM)) in section 2.3. Then, we describe the models that we used to predict the MSW_C, including the model optimization (section 2.4.1), the assessment of the PV importance (section 2.4.2) and the selection of the best model (section 2.4.4). Following, we introduce our waste flow model (WFM), including the calculation of the total amount of MSW (MSW_T, section 2.5.1), plastic waste that enters the environment (P_E, section 2.5.2) and plastic waste that enters the ocean (P_O, section 2.5.3).

2.1 Description of the Data

Here we present all the data we obtained to predict the plastic input to the ocean (compare Fig. 2). We begin with the description of the available data for MSW_C, followed by the description of the PV data used in the MLM to predict the MSW_C. Then we present the data of the waste collection system, the waste composition and the waste disposal, which was used to predict the total amount of plastic waste (P_T) from the MSW_C. This data is summarized in table 2. Finally, we present the geospatial data that we use to predict P_E on a global raster as well as in the WFM and for the ocean current advection model. A detailed overview of the data presented

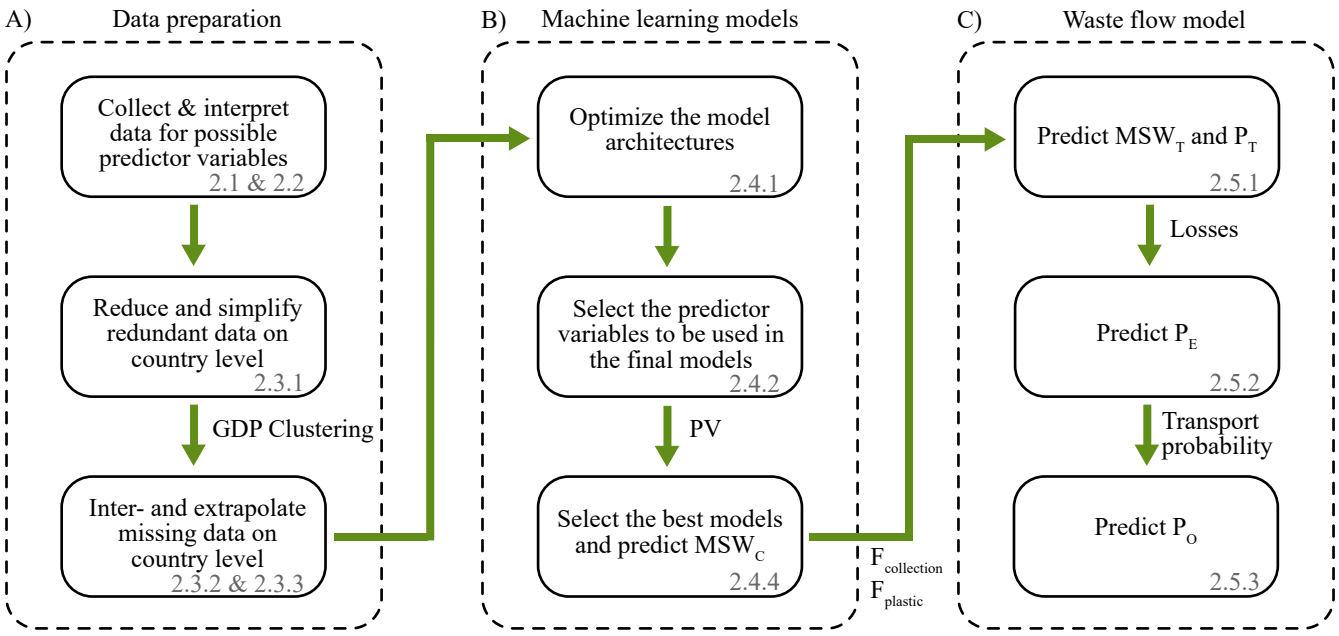

Figure 2. Summary of the processes used to estimate P_O. Each box represents one major step and indicates which section describes this process. The arrows indicate which data or models were used in the process. The column *Data preparation* describes how the data was obtained and modified to be used in the MLM. The column *machine learning models* describes how the MLM were optimized, how they were used to select the optimal PVs and how the the MSW_C was estimated. Finally, the column *waste flow model* describes how we estimated MSW_T and P_T using information on the waste collection and composition data. The WFM used to estimate P_E and P_O is further explained in Figure 8.

in this section can be found in table A.2 in the appendix. This table indicates for each variable how many data points were available per country.

2.1.1 MSW_C data

We obtained data for our target variable MSW_C from the UN database. The dataset includes 1'544 data points for 125 countries in the time span 1990-2015 (compare table 2). No country has a complete dataset for the whole time period (1990-2015, 26 years). However, for four countries (Austria, Luxembourg, Poland and Switzerland) only the data for 2015 is missing. A total of 27 countries has data available for at least 20 years and 64 countries provide data for more than 10 years. 23 countries provide data for from two up to five years and six countries (Angola, Bhutan, Costa Rica, Peru, Sri Lanka and Uganda) provide data for only one year. The data availability is highest in the years 2000-2012 where more than 65 countries have data available.

2.1.2 PV data

We obtained the data for the PVs from the UN database and the World Bank database. The data contains information for 195 countries (officially recognized by the UN) as well as for 36 dependencies and autonomous regions (e.g. Hong Kong, Falkland Islands). For simplicity, these countries, dependencies, and autonomous regions will further on simply be referred to as countries. Table 2 gives a summary of each dataset. Data for the PVs *total population* and *urban population* is available since 1950 and includes projections until 2050. These two datasets cover all 231 countries. The dataset *young population* covers fewer countries (221) and over a shorter time scale from 1960 until 2017. *GDP* covers the time period 1960-2015 and includes data for 202 countries. Data on *greenhouse gas* and *electricity access* was available from 1970-2012 and for 195 ans 203 countries respectively. Data for $F_{recycled}$ is available for 102 countries for the years 1990 and 1995-2015.

2.1.3 Data of waste collection, composition and disposal

We obtained the data on the fraction of population with access to a waste collection system ($F_{collection}$), the fraction of plastic in MSW_C and the fractions of MSW_C that is deposited in landfills ($F_{landfills}$) or dumps (F_{dumps}) from the UN database and from Hoornweg&Bhada-Tata [37]. Data for $F_{landfill}$ is available for 102 countries in the time period 1990-2015. Data for $F_{recycled}$ is only available for 41 countries and in the years 1990 and 1995-2013. Data for $F_{landfills}$ and F_{dumps} is only available for 74 countries in the year 2012. Data for the GNI is available from the World Bank for 204 countries in the time period 1960-2016.

2.1.4 Geospatial data

Population density

We obtained the data on the population density from the Socioeconomic Data and Applications Center (SEDAC). It is available as an ArcInfo GRID file for the years 1990 and 1995 (in a 2.5 arc-minutes resolution) or as a GeoTiff-image for

Table 2. Description and sources of the target variable MSW_C of the MLM, all PVs used in the MLM to predict MSW_C for each country and year and the variables used to estimate MSW_T. *N countries* indicates for how many countries at least one data point was available. *N data points* indicates how many data points were available (including the represented fraction of the maximum possible data points for this time period and number of countries).

Data	Source	Time scale	N countries	N data points	Description
Target variable					
MSW_C	[49]	1990-2015	125	1'544 (48%)	Amount of MSW that was collected in a year.
PVs used in the MLM to predict MSW_C					
Total population	[50]	1950-2050	231	23'331 (100%)	Number of people living in the country.
Urban population	[51]	1950-2050	231	23'331 (100%)	Number of people living in urban areas.
Young population	[52]	1960-2017	195	11'241 (99%)	Fraction of the total population aged 0-14 years.
Electricity access	[53]	1990-2014	204	4'931 (83%)	Fraction of households with access to electricity.
GDP	[54]	1960-2016	202	8'913 (77%)	Gross domestic product per country in USD, current conversion rate.
Energy consumption	[55]	1990-2015	173	3'644 (78%)	Energy consumption per person (in kg oil equivalents).
Greenhouse gas	[56]	1970-2012	199	8'060 (94%)	Total emissions of greenhouse gases in CO_2 equivalents.
$F_{recycled}$	[57]	1990, 1995-2013	78	916 (59%)	Fraction of MSW_C that is recycled.
Variables used in to estimate MSW_T					
GNI	[58]	1960-2016	204	8'901 (70%)	Gross national income per country in USD, current conversion rate.
$F_{collection}$	[59]	1990-2015	102	958 (36%)	Fraction of MSW_T that is collected at the households.
$F_{plastic}$	[37]	2012	103	103	Fraction of plastic in MSW_C.
$F_{landfills}$, F_{dumps}	[37]	2012	74	74	Fraction of MSW_C that ends in either landfills or dumps.

the years 2000, 2005, 2010 and 2015 (in a 30 arc-seconds resolution) (table 3). We processed the data to create data of the same grid size and with continuous information for the years 1990-2015. To this end, the data was loaded into RStudio using the *raster* function from the *rgdal* package [60] and then converted to a 0.1x0.1° raster (6 arc-minutes) using the *aggregate* function from the *raster* package [61]. The value of each aggregated cell was calculated using the mean of all original cells. Some aggregated cells (at the shore line) of the 2.5 arc-minute rasters were therefore outside the scope of the target raster (e.g. ocean cells). The aggregated cells were compared to the rasterized coastline and cells that laid outside the target raster (e.g. in the ocean) were dropped. To create continuous data, the population was then interpolated for the missing years. This was done using the growth rate of each country's population as obtained from the predictor variable *total population*. However, the urbanization in this time period was neglected and the population thus increases linearly in every cell.

River shape

We obtained a shape file including 342 river centrelines from NaturalEarth (table 3). This data is used in the WFM and includes all major rivers at the global scale. Lakes are represented with their centrelines and not with their full area. Intermittent rivers (large, seasonal rivers) are not included in this dataset. We created a raster of all rivers using the *rasterize* function of the *raster* package [61] for further use in the WFM.

Ocean currents

We obtained the ocean currents data from the GlobCurrent project [62]. The dataset includes daily averaged satellite-based estimates of the surface Eulerian total current from 1993 to 2017 on a 0.25x0.25° global grid (table 3). The Eulerian total current, representing the sum of geostrophic and Ekman current components, was chosen over the simpler pure geostrophic or Ekman current datasets as it was found to represent drifting particle distributions more realistic [62]. However, a known issue of this dataset is that current velocities in local areas close to the coast can be erroneous [62]. Furthermore, the Mediterranean is not included in the dataset and particles released there will thus not be advected. A graphical representation of the data can be seen in Figure 3.

In this section we described the availability and completeness of the individual datasets. In the following section we present the correlations among the datasets.

2.2 Trends and Correlations among the Predictor Variables and with the Target Variable

In this section we present trends and correlations in the datasets that were described in the previous section. We first describe the correlations among the PVs and our target variable and compare this to our expectations as described in the introduction. Then we show correlation among the PVs themselves.

Correlations among MSW_C and the PVs

Here we present the Pearson and Spearman correlations of MSW_C with the PVs (Fig. 4). The Pearson correlation measures the relationship between linearly related variables while the Spearman correlation can be used to assess monotonic or exponential relationships.

The Pearson correlation (Fig. 4 A) of the PVs *energy consumption* (0.52) and *GDP* (0.52) are >0.5. *Young population* (-0.31) is the only negative correlation. $F_{recycled}$ (0.24) has the correlation closest to zero. The Spearman correlation (Fig. 4 B) shows similar values as the Person correlation. Here, *electricity access* is the only PV whose difference to the Pearson correlation is >0.07. This can be explained by looking at the scatterplot of *electricity access* and MSW_C in Figure A.1 in the appendix. It can be seen that *electricity access* the data is not linearly distributed and shows an exponential trend which is captured by the Spearman correlation.

Correlations among the PVs

The Pearson correlation (Fig. 4 A) reveal that the correlations among the PVs cover a wider range (0.18 - 0.74) than with the target variable MSW_C. Correlations >0.5 can be found between *young population* and *electricity access* (-0.74), between *greenhouse gas* and *energy consumption* (0.61), between *energy consumption* and *urban population* (0.57) and between *urban population* and *electricity access* (0.55). We expect that if two variables are highly correlated, one will be removed in the model optimization process (described in section 2.4.2). Highly correlated variables include similar relationships towards MSW_C thus potentially redundant information.

The Spearman correlation (Fig. 4 B) among the PV cover a slightly smaller range (0.17 - 0.63) than the Pearson correlations (0-18 - 0.74) and a slightly larger range than with MSW_C (0.27 - 0.58). Compared to the Pearson correlations, *electricity access* shows a stronger Spearman correlation for all PVs except *urban population* and *young population*. When looking at the scatterplots of these PVs A.2, a linear relationship can be seen, which explains the higher Pearson correlation. Furthermore, the Spearman correlations between *GDP* and *greenhouse gas* (0.43) and between *urban population* and *young population* (0.17) are lower respectively higher than the Pearson correlations (0.29 and -0.44 respectively). The difference in the Spearman and Pearson correlations among the other PVs is <0.15.

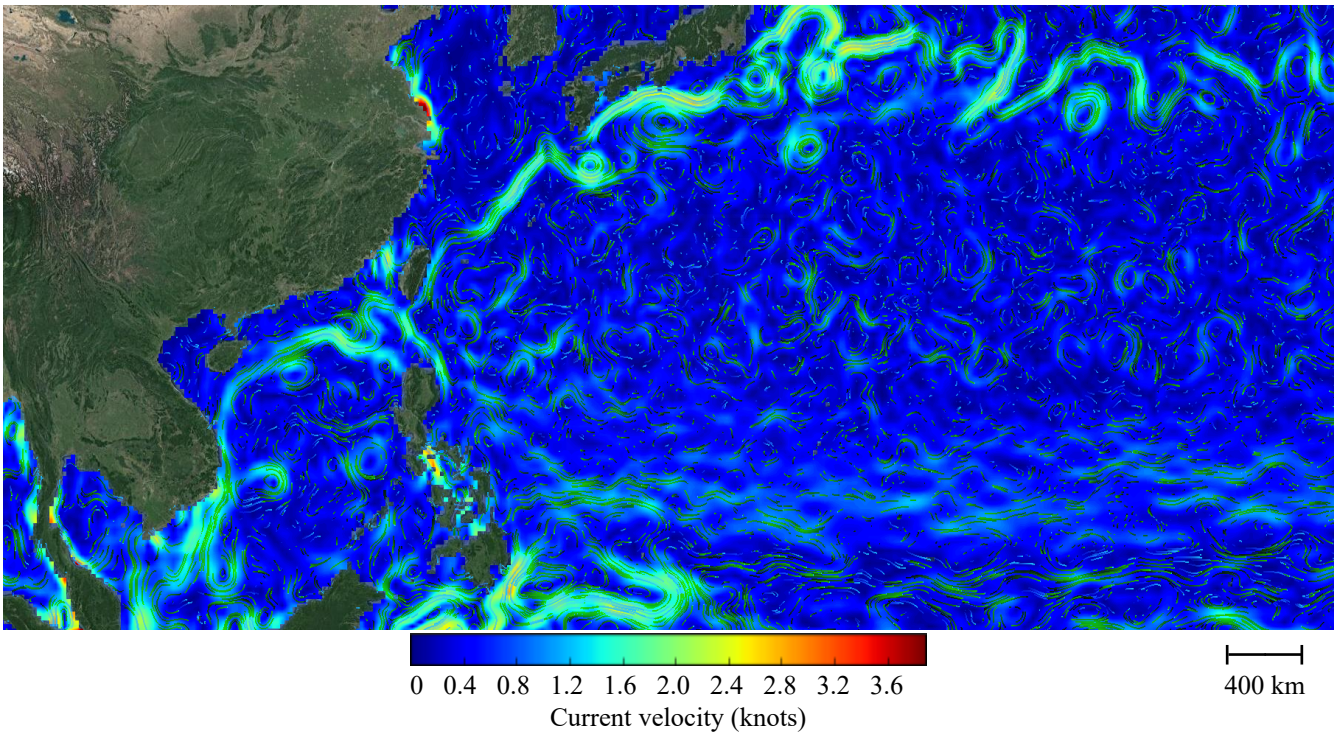

Figure 3. Visualization of the GlobCurrent surface Eulerian total current data in the South China Sea and the western North Pacific as used in the PARCELS simulations, including the current velocity as the background layer and indicating current streamlines in green. Visualization created using the Syntool visualization website and manually edited [63].

Table 3. Description and sources of the geospatial data. The population density is used in the prediction of MSW_C on a 0.1x0.1° raster. The river dataset is used in the calculation of plastic transport to the oceans. The ocean current data is used to simulate the advection of plastic particles in the ocean.

Data	Source	Time scale	Type	Description
Population density	[64]	1990, 1995	Raster with a resolution of 2.5 arc-minutes	Population density.
Population density	[65]	2000, 2005, 2010, 2015	Raster with a resolution of 30 arc-seconds	Population density.
Rivers	[66]		Shape file of lake and river centres with a resolution of $1{:}10^6$	Polyline of major river and lake centres.
Ocean currents	[67]	1993-2017	NetCDF file with a resolution of 0.25x0.25°.	Total Eulerian currents (sum of geostrophic and Ekman components).

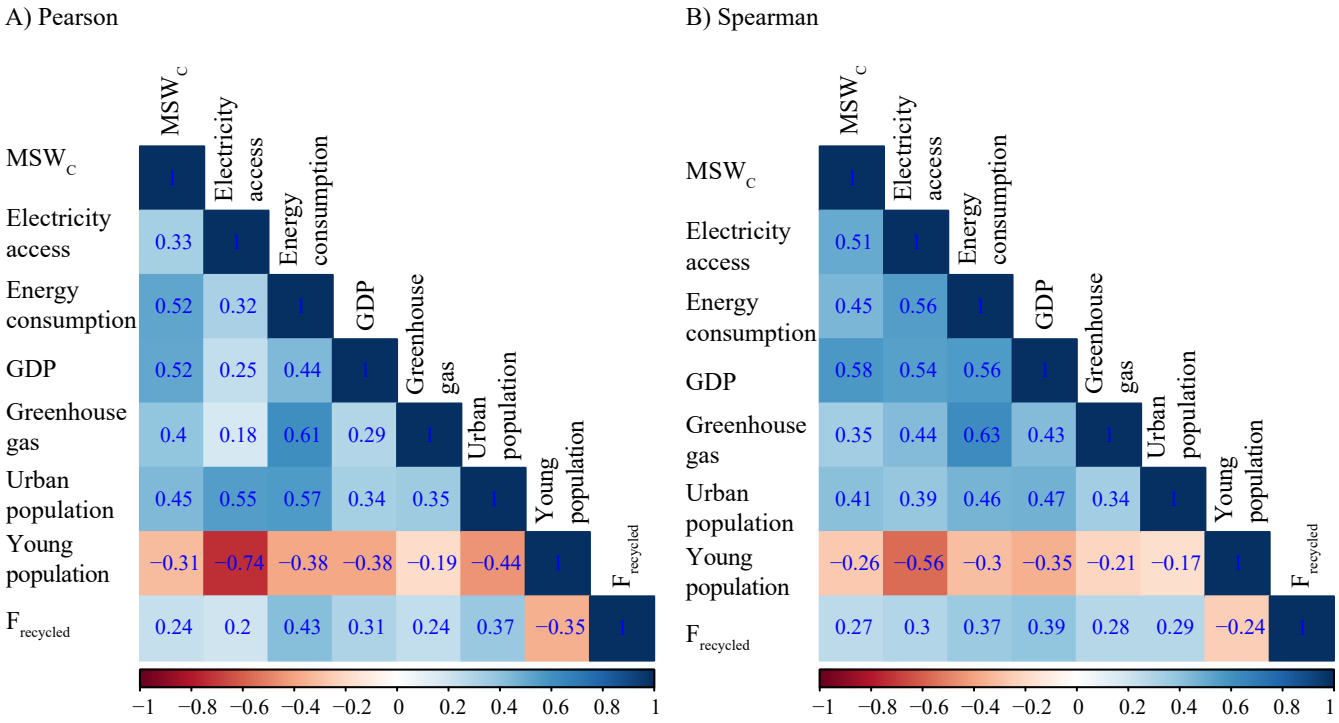

Figure 4. Pearson and Spearman correlations between MSW_C and the predictor variables as well as among the predictor variables. We expect that highly correlated variables are removed in the variable selection procedure explained in section 2.4.2.

To further visualize the relationships between the PVs, we created a dendrogram that uses the correlation among the PVs as measure of similarity (Fig. A.3 in the appendix). If can be seen that the PVs *electricity access* and *young population* as well as *energy consumption* and *greenhouse gas* are grouped in both the Pearson and Spearman dendrogram. Additionally, *GDP* and MSW_C are also grouped. Again, we expect that, if two PVs are grouped, one will be removed in the model optimization process.
In the next section we describe how we prepared the data for the use in the MLM.

2.3 Data Preparation for the MLM

The obtained data as described above is often incomplete regarding the number of countries or years that are represented as well as the number of data points (compare table 2). As the MLM and the WFM need complete PV datasets, the data had to be prepared for the modelling steps (compare Fig. 2). We first reduced and simplified redundant data and then inter-and extrapolated missing data. This increased the data completeness by filling the gaps. Following, we rasterized the data, which is needed for the geospatial analyses in the WFM.

2.3.1 Reduction and simplification of redundant data

Here we describe how we reduced and simplified the datasets to increase the data completeness. Following measures were performed:

- We grouped small islands states in Oceania (e.g. Federated States of Micronesia, Marshall Islands, N=16) as well as in the Caribbean (e.g. Jamaica, Cayman Islands, N=17) and applied the mean values (compare table 4). We thus assume that they are similar in their MSW production rates.
- We dropped countries that did not provide GDP data (e.g. Niger, North Korea). This includes 28 countries as presented in table 4.
- We cropped all datasets to the years 1990-2015, as these years contain substantially more data than in the period 1960-1989. Data for 2016 and 2017 was often not yet published.

These measures reduced the number of countries that we used to model the MSW_C from 231 to 173. The dropped countries represent 0.78% of the world's population. Table 4 gives an overview of the dropped and grouped countries. Together with the reduction of the time period, we reduced the number of needed data points from 15'708 (231 countries and 68 years from 1950-2017) to 4'498 (173 countries and 26 years from 1990-2015). This increased the data completeness of the modified datasets (compare table A.1). As the MLM require complete datasets for the PVs, and these measures did not achieve 100% data completeness, we had to estimate the missing data.
We found two cases of missing data. First, countries that provided PV values for several years but the data does not cover the whole time period. This missing data was inter or extrapolated as described in the next subsection. Secondly, countries that did not provide data for a PV. In this case, the

data was estimated based on similar countries. This process is described section 2.3.3.

2.3.2 Inter- and extrapolation of partially missing data

To estimate the missing data of countries where more than 2 data points were available in the years 1990-2015, we calculated the missing values using either a linear or polynomial fit. A linear fit was used for the datasets *electricity access*, *energy consumption*, *greenhouse gas*, $F_{collection}$ and $F_{recycled}$ as these PVs show a linear trend in the available data (compare table A.1). We used a 2nd order polynomial fit for the GDP as the available data suggests a non-linear trend (compare table A.1). We did not allow the calculated data from the linear fit of $F_{collection}$ to fall below the respective continental minimum as presented by Hoornweg&Bhada-Tata [37] (applied values for Africa: 18%, Asia: 50% Europe: 60% and South America: 11%). Missing data in the the datasets F_{dumps} and $F_{landfills}$ was replaced by the mean values according to the countries income class (high-income: >12'476 USD/year, upper middle-income: 4'046-12'475 USD/year, lower middle-income: 1'026-4'035 USD/year and low-income: <1'025 USD/year) as presented in Hoornweg&Bhada-Tata [37].

2.3.3 Estimation for countries with completely missing data

To get an estimate of each variable for countries where no data was available to extrapolate (compare tables A.1 and A.2), we replaced the missing data with the mean from countries with a similar GDP. Therefore, we sorted all countries into clusters according to their similarity in their GDP trajectory (Fig. 5). Countries where no data was available to run a linear or exponential fit were given the mean values of the corresponding cluster (compare table A.1). This was performed on the datasets *electricity access*, *energy consumption*, *greenhouse gas*, $F_{collection}$, $F_{plastic}$ and $F_{recycled}$. The clustering was done using the *kmlShape* package [68] version 0.9.5. This package calculates K-means clusters for longitudinal data using shape-respecting distance and shape-respecting means. The ideal number of clusters was determined by minimizing the sum of differences between the trajectory of each country and the respective cluster trajectory mean, running a maximum of 15 clusters (compare Fig. A.4 in the appendix). The optimum result was obtained by creating 7 clusters and is presented in Figure 5. It can be seen that the clusters 1-6 include countries from at least two continents and only the cluster 7 include solely the European countries Liechtenstein and Monaco (compare also table A.2). Table A.1 indicates the number of counties for which the data was estimated using this approach.
The now complete datasets are ready to be used in the MLM, but the geospatial analyses in the WFM require that we rasterize the data. This is described in the next section.

2.3.4 Rasterization of Country Level Data

After performing the steps described above, the datasets presented in table 2 now have a data completeness of 100% and are ready to be used in the MLM. For the WFM, the data needs to be rasterized to allow geospatial analyses. Here we describe how we created a 0.1x0.1° raster for each dataset as presented in table 2 for each year in the period 1990-2015.
We obtained a spatial polygon file of the administrative boundaries for each country using the *getData* function from the *raster* package [61]. The data of each dataset was assigned to each cell of the 0.1x0.1° raster whose cell-centre lays in the corresponding spatial polygon. Due to its small area of just 2 km^2, the country of Monaco is not represented in any raster as no cell has its centre inside Monaco's boundaries. Thus, the data of Monaco was added manually to the cell containing the city centre of Monaco in each raster. Therefore, the country France "lost" a cell to Monaco. This process was however not needed for other city states or micro states such as Singapore or Kuwait.

With this, the data is now prepared to be used in the MLM and WFM. In the next section we describe how we used the MLM to predict MSW_C using a selection of PVs and section 2.5 describes the WFM.

2.4 Modelling MSW_C

After describing the acquisition and preparation of the PVs we now present the modelling approach used to predict MSW_C. As the observed trends between the PVs and MSW_C are not always linear (compare section 2.2) we chose to use both NNET and RF models to predict MSW_C. To avoid overfitting, we run tests to find the best compromise between the least complex architecture, a minimal amount of PVs and a minimal error in the predictions. Less complex models do not only prevent overfitting but were also found to be better in predicting unseen data [69]. Furthermore we asses the stability of the models to the size of the training data and biases in the training data. We use two different models, including a neural network (NNET) and a random forest (RF) which allows a more robust estimate as the model results can be averaged. Both models use the available MSW_C values as target variable and the PVs as described in section 2.3 and table 2. We first describe how we optimized the architecture of the models. Following, we describe how we selected the best PVs and then show how we assessed the model stability. In the end of this section we describe how we selected the best NNET and RF model that will then be used to predict MSW_C.

2.4.1 Model architecture optimization

We used a individual procedure for the NNET and RF model to optimize the architectures. In both cases we used all points where MSW_C data was available (N=1'394). We selected the corresponding values of the PVs and normalized them. The target variable MSW_C was also normalized. This was done to prevent a bias in the models [69]. We then performed following steps:

Architecture optimization of the NNET

The NNET was created using the R package *neuralnet* [70] version 1.33. We tested several network architectures by train-

Table 4. Countries, dependencies and autonomous regions that were dropped or grouped from the analysis. Dropped countries did not provide GDP data which was needed for creating clusters that allow the estimation of missing data (section 2.3.2). We grouped the Caribbean and Oceania islands because we assume that they behave similarly regarding the MSW_C production.

Dropped due to missing data					
Anguilla	Caribbean Netherlands	Cook Islands	Curaçao	Falkland Islands	French Guiana
Gibraltar	Guadeloupe	Holy See	Kyrgyzstan	Martinique	Mayotte
Montserrat	Niger	Niue	North Korea	Palestine	Réunion
Saint Helena	Saint Martin	Saint Pierre and Miquelon	Sint Maarten	Tokelau	Turks and Caicos Islands
Tuvalu	United states Virgin Islands	Wallis and Futuna Islands	Western Sahara		
Grouped to "Caribbean Islands"					
Antigua and Barbuda	Aruba	Bahamas	Barbados	British Virgin Islands	Cayman Islands
Cuba	Dominica	Dominican Republic	Grenada	Haiti	Jamaica
Puerto Rico	Saint Kitts and Nevis	Saint Lucia	Saint Vincent and the Grenadines	Trinidad and Tobago	
Grouped to "Oceania Islands"					
American Samoa	Federated States of Micronesia	Fiji	French Polynesia	Guam	Kiribati
Marshall Islands	Nauru	New Caledonia	Northern Mariana Islands	Palau	Papua New Guinea
Samoa	Solomon Islands	Tonga	Vanuatu		

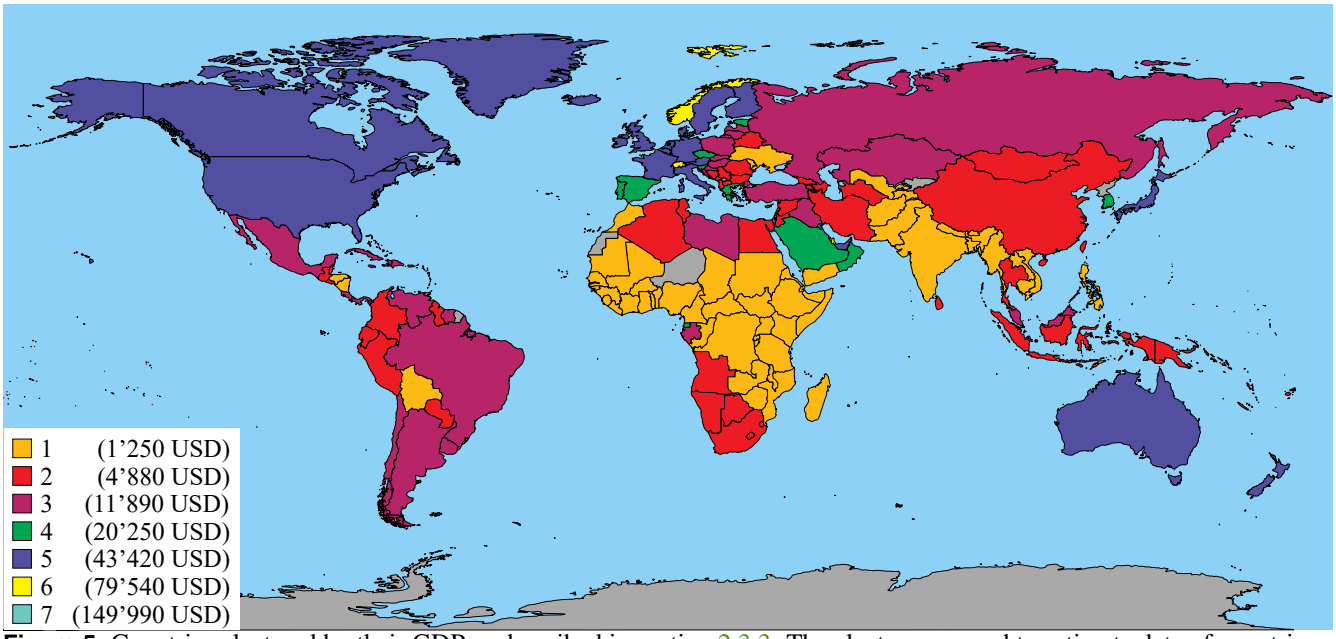

Figure 5. Countries clustered by their GDP as described in section 2.3.3. The clusters are used to estimate data of countries that did not provide any data and did thus not allow inter-and extrapolation. The numbers in brackets indicate the average GDP per cluster. Grey countries are not included in the calculations due to missing GDP values (compare the tables 4 and A.2). Note that cluster 7 includes only the countries Monaco and Lichtenstein and is thus not visible on the global map. Table A.2 gives detailed information on the GDP and cluster of each country.

ing ten models on the same, randomly chosen 90% of the scaled data. To evaluate the model we performed a 10-fold cross validation (CV) and calculated the mean squared error (MSE). This process was repeated ten times on newly randomized data samples to avoid a bias in the training data. Thus, a total of 100 individual models were created. We retained the model with the lowest MSE and compared it to the best model of the other model architectures (Fig. 6). Our optimal architecture for the NNET model consists of one layer with five nodes (MSE:0.03). While adding more nodes to the layer does not decrease the MSE, removing nodes results in a loss of accuracy. While some architectures consisting of 2 layers (e.g. three and two nodes per layer, Fig. A.5) were equally successful in prediction, we preferred the one layer model as to limit the model complexity. A summary of all layer and node combinations can be found in Figure A.5 in the appendix.

Architecture optimization of the RF

The RF was created using the R package *randomForest* [71] version 4.6-14. The RF creates a "forest" of 500 unique decision trees and averages the outcome to obtain the best prediction. We chose 500 trees as increasing this number did not decrease the MSE (Fig. A.6) [72]. Each tree has a random element in its creation and only selects a subset of the PVs in each branch. We optimized the size of this subset by using the *tuneRF* function, which calculates the out of bag error for each try [71]. We run the *tuneRF* function 100 times using different randomly sampled 90% of the data to avoid a bias in the training data [72]. For our data, a *mtry* value of two results in the smallest out of bag error in 96% of the tries while in 4% using three variables is preferred. To reduce complexity and prevent overfitting [72], we thus set *mtry* to two.

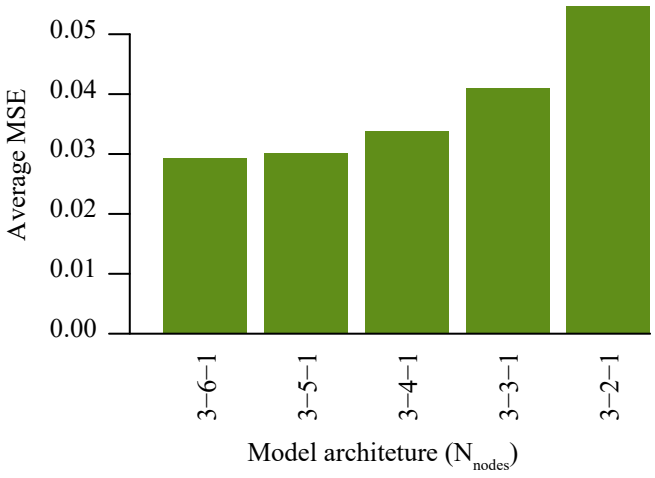

Figure 6. Summary of the average MSE for the different one layer model architectures for the NNET. The first number indicates the amount of PVs, the second number shows the number of nodes in the hidden layer and the third number stands for the number of outputs. To avoid overfitting of the NNET we chose the simplest architecture that has a low MSE.

2.4.2 Predictor variable selection

Besides using a simple architecture, using a minimal number of PV can help to prevent overfitting [69, 72]. We ran several tests to select which PVs are most important in the model creation. We use the following techniques to assess the importance of each PV: removal, random permutation and univariate runs. Following we describe these procedures:

Removal: We run both MLM while sequentially removing one parameter at a time. By removing one parameter from the training data we assess the importance of this variable in the model creation [69]. For each removed PV, we trained ten models on the same randomly sampled 90% of the data. This was repeated nine more times using newly sampled training data to avoid a bias in the training data. Thus, a total of 100 models were created. We then remove the parameter that has the least effect on the model performance (MSE). This process was repeated with the remaining PVs until only one PV was left.

Random permutation: Analogue to the removal process we run both models while randomly permuting one parameter at a time. This means that for one PV, the values were randomly replaced with values from within the same PV. If the PV is important in the model creation this is expected to create larger errors than for less important PVs [72]. For each randomly permuted PV, we trained ten models on the same randomly sampled 90% of the data. This was repeated nine more times using newly sampled training data to avoid a bias in the training data. We then remove the parameter that has the least effect on the model performance (MSE). This process was repeated with the remaining PVs until only one PV was left.

Univariate: We ran the model using only one PV at a time and calculated the MSE. Here, we expect that PVs that are important for the model creation have lower MSE values [69].

We then compared the effect of the removal, random permutation and univariate runs on the MSE. Figure 7 shows the example of PV removal in the RF model. As soon as removing a PV created a big increase in the MSE, we assumed that this PV, together with the remaining PVs, is crucial in the model creation [69]. In this case, removing the PVs, $F_{recycled}$, *GDP*, *energy consumption* or *electricity access* has only a small effect on the MSE. If *young population* is removed however, the effect on the MSE is stronger. Thus, we selected the three PVs *young population*, *greenhouse gas* and *urban population* as essential for the model.

A summary of the PV selection (including both models and all selection processes as described above) can be seen in table 5. While no PV was always the most or least important, clear trends can be seen. The PVs *urban population* and *greenhouse gas* are the most important in both the average rank (1.2 and 2.2 respectively) as well as the median rank (1 and 2 respectively). These two PVs were always among the three most important PVs throughout all selection types. The PVs *young population* and *energy consumption* have an average rank of 4.0, but the median of latter is higher (3.5

and 5 respectively). The parameter removal and parameter shuffling processes both favour *young population* and only the univariate selection process favours *energy consumption*. Thus, *young population* was preferred. $F_{recycled}$ was usually unimportant and thus has the lowest average and median rank (6.5 in both). As a result of these tests we selected *young population*, *greenhouse gas* and *urban population* as our most important PVs. They were used in the prediction of MSW_C while the others were dropped.

2.4.3 Testing the model stability by varying the training data

After optimizing the model architectures and selecting the PVs, we were interested in the stability of the models with respect to the size of the training data and biases in the training data. Therefore, we separately trained 100 models using all the available data from each continent, cluster or GNI classification as biased training data. We selected the model with the lowest MSE to compare the effect of the selected bias in the training data. Additionally we compared the results of the training data size from models using randomly selected 90%,80%,...,10% of the available data. These tests allow us to assess the robustness of the MLM to uncertainty and loss of the underlying data as well as the dependence of the MLM results on PV data substructures. The results of the stability tests are presented in section 3.6.

2.4.4 Selection of the best model

After optimizing the model architectures and selecting the PVs, we wanted to select the NNET and RF model that yields the best results. We therefore trained 10'000 NNET and RF models on randomly selected 90% of the available data. We then used several techniques to assess the prediction performance, which will now be presented:

CV performance: In a first step, each generated model is retained or excluded based on their performance in the CV. The CV was done while creating the model (section 2.4.1). We only retain models whose MSE of the prediction using the unseen test-dataset is smaller than the MSE of at least one prediction using one of the nine training-datasets of the CV. This drops models that were not able to predict unseen data within the same MSE range as the training data used in the CV.

MSE: In a second step, we rank the models by their MSE of the prediction of the test data. This selects models that were better in predicting the unseen data.

Errors for selected countries: In a third step we ranked the models according to their prediction performance for selected countries. This is done to avoid large errors for countries with a big global impact. We ordered the models by their performance on the 20 countries that contribute the most to P_O according to Jambeck et al. [4], and for which we had data available. This includes the eleven countries China, Indonesia, the Philippines, Sri Lanka, Egypt, Malaysia, Algeria, Turkey, Brazil, Morocco and the United States of America. The model prediction error of these countries was multiplied with the country's marine plastic input according to Jambeck et al. [4]. This prefers models that show low prediction errors for the countries with the highest influence on P_O.

We then selected one NNET and RF model each. The selected model has one of the smallest MSE and its error for the prediction for the top 20 countries is also among the smallest. Additionally, its overall prediction error for all countries is as close to the smallest prediction error of all models. The selected NNET and RF models were finally used to predict the MSW_C for each country and year. We present these models in section 3.1.

The above described NNET and RF models are used to predict the MSW_C for each of our 173 countries for the years 1990-2015. In order to create geospatial estimates in the WFM, the MSW_C data has to be rasterized. This was done as described in section 2.3.4. The procedure to estimate P_E and P_O from MSW_C are described in the next section.

2.5 Waste Flow Model

By running the MLM we estimated MSW_C for each country in the time period 1990-2015. Next, we describe how we calculated how much plastic enters the ocean (P_O) on a 0.1x0.1° raster using our WFM. First we present the calculation of the total amount of produced MSW (MSW_T), consisting of the MSW inside the waste collection system (MSW_C) and uncollected waste. Using this information, we then calculated the total amount of plastic waste (P_T), again inside and outside the waste collection system. These steps are described in section 2.5.1. Then, we show how we estimated the amount of plastic that is entering the environment (P_E) in section 2.5.2. This includes both plastic lost from the waste collection system and littering as well as from uncollected plastic waste. Finally, we present how we estimated the amount of P_E that enters the ocean by using a distance based transport probability (section 2.5.3). This process is also visualized in Figure 8.

2.5.1 Calculating the total amount of MSW and the total amount of plastic waste

MSW_C, as calculated by the MLM using the PVs, only includes the MSW that was collected in each year. It does not include MSW that was not collected at the households (e.g. due to a missing collection system or due to missing regulation on MSW treatment [37]). To get an estimate of how much waste is uncollected, we used the fraction of population with access to the waste collection system ($F_{collection}$). The total amount of MSW (MSW_T) was thus obtained by adding the uncollected waste (outside waste treatment system) to MSW_C (inside waste treatment system). By further applying each country's fraction of plastic in MSW ($F_{plastic}$) we calculated the total amount of plastic waste produced (P_T) on a 0.1x0.1° raster (both inside and outside the waste treatment system). In the next section we describe how this was used to estimate P_E.

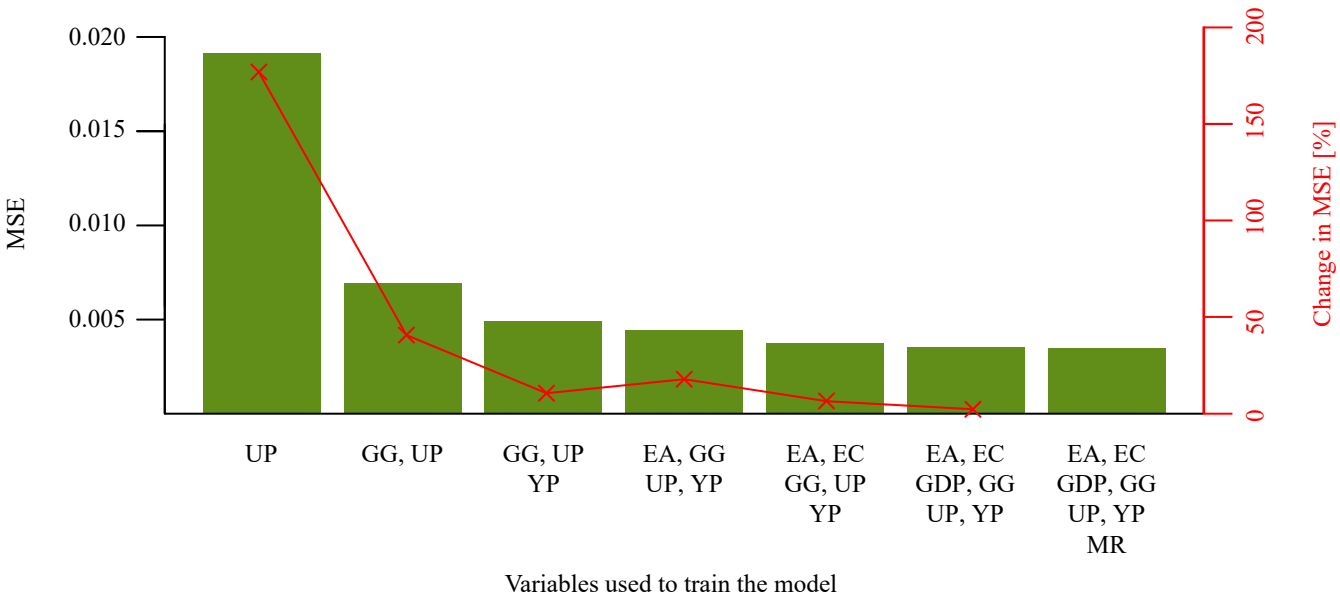

Figure 7. Results of the PV selection by parameter removal for the RF model. The bars indicate the average MSE of all 100 RF with the PVs indicated. The line indicates the change in MSE between two bars. A bend in the line represents a strong decrease in the model accuracy which indicates that a PV crucial for the model was removed. For this example, the PVs GG, UP and YT were selected (compare table 5). Abbreviations: EA = electricity access, EC = energy consumption, GG = greenhouse gas, UP = urban population, YP = young population and MR = $MSW_{recycled}$.

Table 5. Results of the PV selection. The numbers indicate the rank of the PV in the certain process (e.g. urban population was always the most important PV). A lower rank indicates a higher importance compared to high-ranked PVs. The PVs in bold were selected for the final model (compare Fig. 7).

PV	Parameter removal		Parameter shuffling		Univariate		Average	Median
	NNET	RF	NNET	RF	NNET	RF	Rank	Rank
Electricity access	4	4	7	4	5	7	5.2	4.5
Energy consumption	5	5	5	5	1	3	4.0	5
GDP	6	6	3	7	4	4	5.0	5
Greenhouse gas	**2**	**2**	**2**	**2**	**3**	**2**	**2.2**	**2**
Urban population	**1**	**1**	**1**	**1**	**2**	**1**	**1.2**	**1**
Young population	**3**	**3**	**4**	**3**	**6**	**5**	**4.0**	**3.5**
MSW recycled	7	7	6	6	7	6	6.5	6.5

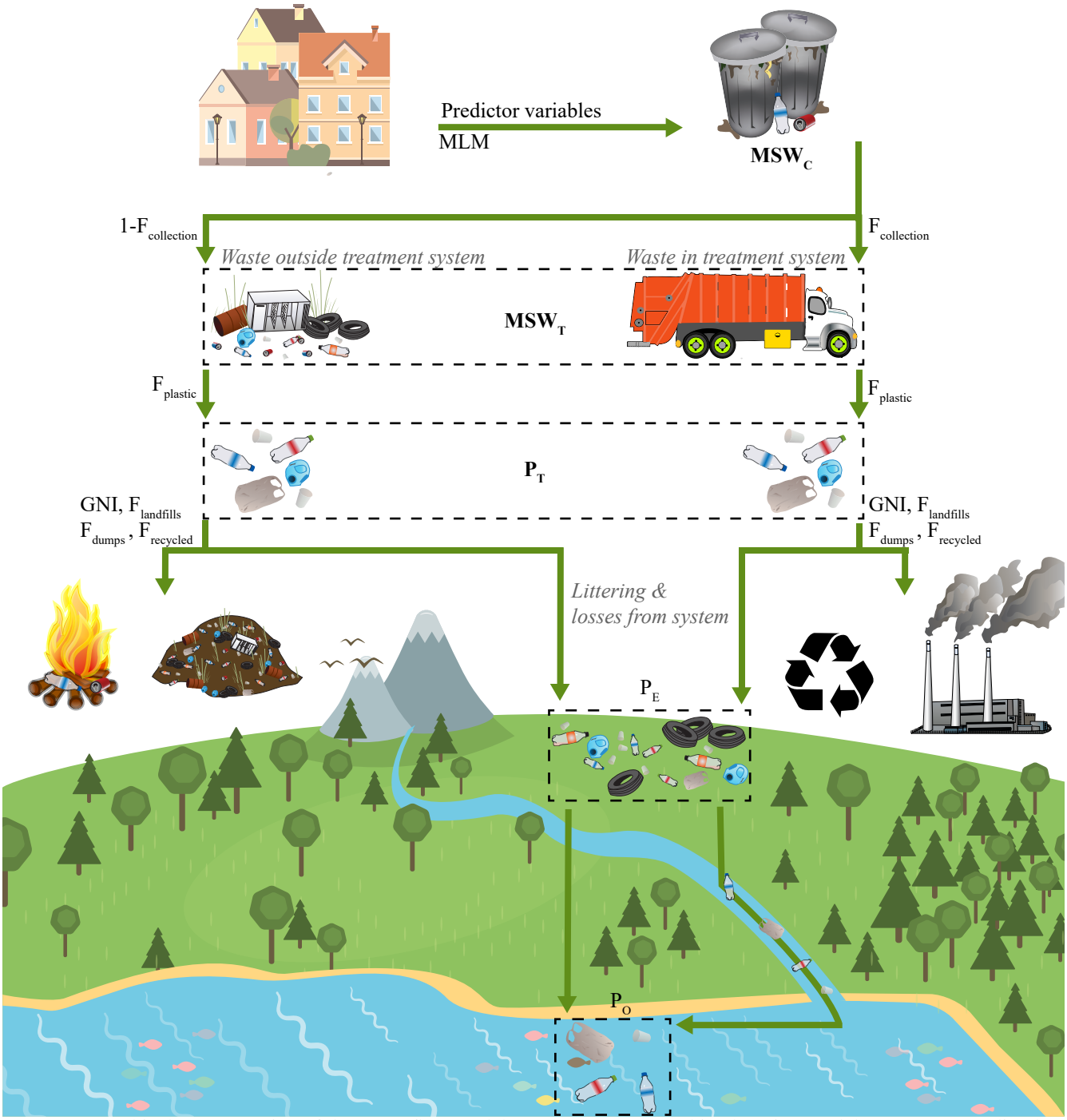

Figure 8. Graphical representation of the waste flow model (WFM). We use machine learning models (MLM) to predict the amount of municipal solid waste collected at the households (MSW_C). This process is described in section 2.4. Using the fraction of population with access to the waste collection system ($F_{collection}$) we estimate the amount of uncollected MSW. The sum of MSW_C and uncollected MSW represents the total amount of MSW (MSW_T) that is produced. This is described in section 2.5.1. We then use the fraction of plastic ($F_{plastic}$) to estimate the amount of plastic inside and outside the waste collection system (P_T). Then, we predict the losses of P_T to the environment (P_E). The losses are dependant on each countries waste collection system as described by the fractions of waste treated by recycling ($F_{recycled}$), in landfills ($F_{landfills}$) or in dumps (F_{dumps}). A more detailed description can be found in section 2.5.2. Finally, we estimate the amount of P_E that reaches the ocean (P_O). This is done by applying distance based probabilities as described in section 2.5.3.

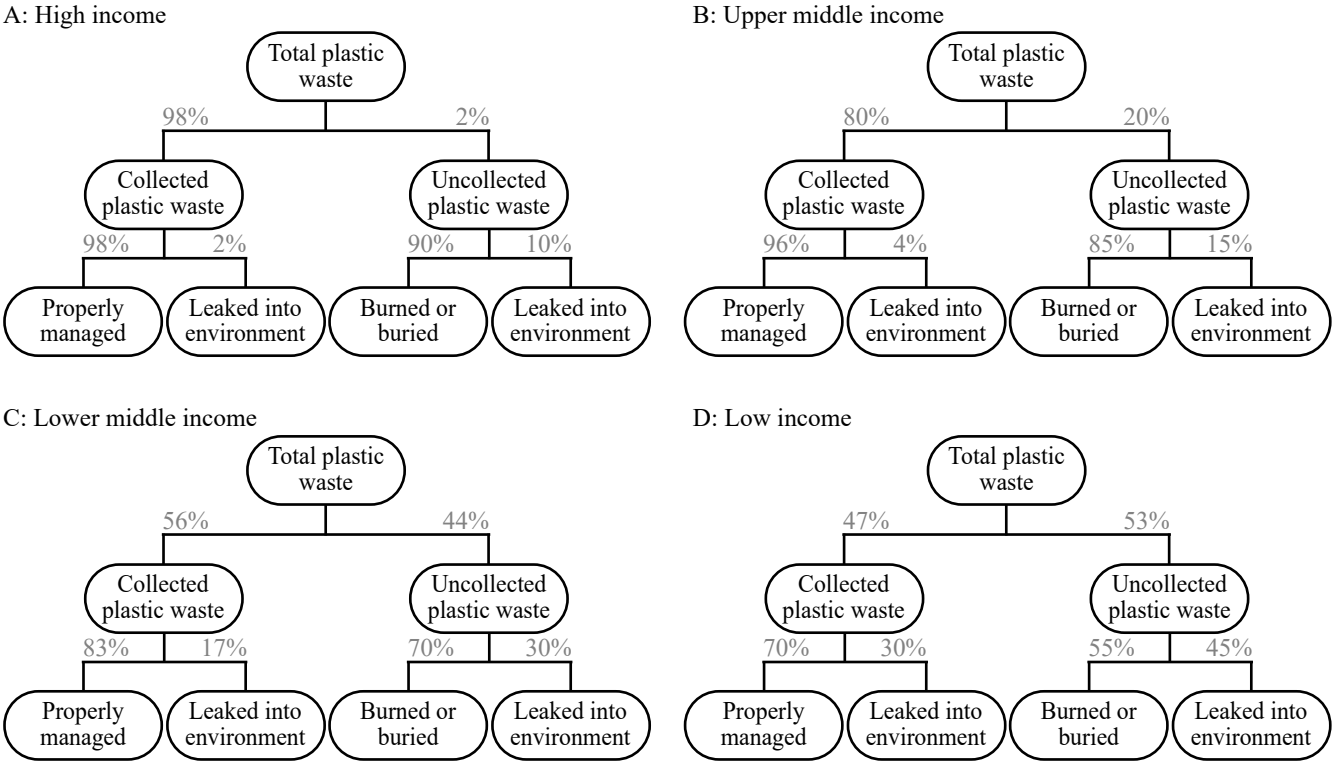

Figure 9. Waste flow models for high-income (A), upper middle-income (B), lower middle-income (C) and low-income (D) countries. The fractions of collected and uncollected plastic waste are the average of $MSW_{collection}$ in 2015 and the country specific values were considered in the model. The fractions of leakage in B and C represent values of China and the Philippines respectively and are taken from the OceanConservancy and the McKinsey Center for Business and Environment [73]. Fractions of leakage for A and D are assumptions. In A, on average 2.2% of all produced plastic waste leaks into the environment while in B, C and D averages of 6.2%, 22.7% and 38.0% are leaked respectively.

2.5.2 Calculating the amount of plastic entering the natural environment

To estimate how much of the plastic inside and outside the waste collection system enters the environment, we needed to account for the MSW treatment situations of each country. As represented in Figure 8, collected plastic waste can either be recycled, incinerated or it can be deposited in both landfills (government operated and controlled) or dumps (sites that are not under government control) [37]. Uncollected plastic on the other hand can be burned locally or put into dumps [37]. Losses from uncollected plastic to the environment are more likely to occur and a study in China, Indonesia, the Philippines, Thailand and Vietnam concluded that approximately 75% of plastic in the environment comes from uncollected plastic waste [73] while the remaining 25% leaked from the waste collection system. We therefore had to treat collected and uncollected waste differently to estimate the losses to the environment. We assume that the flux from P_T to P_E from both inside and outside the waste treatment system is dependant on the countries income [37]. Thus, we treated each income class differently, which is represented in Figure 9 A-D. Following we present how the fluxes from P_T to P_E were estimated for plastic inside and outside the waste treatment system.
Unfortunately it was not possible to find precise information on the loss of collected and uncollected plastic. However, the study by OceanConservancy and the McKinsey Center for Business and Environment [73] investigated plastic leakage fractions in the Philippines and China in detail. They found that 84% of plastic that enters the environment in China comes from uncollected MSW (Fig. 9 B) while 74% of leaked plastic in the Philippines comes from waste inside the collection system (Fig. 9 C). To adapt these values to the other countries, each country was grouped based on its GNI (according to the world bank definition [37]): high-income (Fig. 9 A), upper middle-income (Fig. 9 B), lower middle-income (Fig. 9 C) and low-income (Fig. 9 D). The loss fraction of collected and uncollected plastic of lower middle-income countries was then calculated by multiplying the reported fraction of loss of the Philippines (17% for collected plastic, 31% for uncollected plastic ,Fig. 9 C)) with a scaling factor. This scaling factor was derived by dividing the fraction of dumps and landfills of the country by the fraction of dumps and landfills in the Philippines (compare table A.2). Thus, lower middle-income countries that dispose more of their waste in landfills and dumps than the Philippines are assumed to have a loss fraction that is bigger than the Philippines. Upper middle-income countries were compared to China in the same way. However, for high and low-income countries, no data could be obtained. Thus, the assumption was made that high-income countries loose 2% of their collected waste (littering rate used in Jambeck et al. [4]) and 10% of their uncollected waste (Fig. 9 A), while low-income countries loose 30% and 45% respectively (Fig. 9 D).
With the estimates of P_E we then predicted P_O using a distance based probability. This is described in the next section.

2.5.3 Probability of plastic waste transport

In order to estimate P_O based on P_E, we applied a distance based probability of transport. This includes riverine transport as well as a land-based transport to the shores. These two types of transport will be explained in this section.

Land based transport
A raster containing the minimum distance of each land cell to the closest ocean cell was calculated using the *distance* function from the *raster* package [61]. The maximum travel distance was set to 100 km and the probability of mismanaged plastic waste to reach the ocean was set to 0% outside this range. Cells inside the range obtained a probability that rises proportionally to the distance from 0% at the maximum distance to 100% at the shore ($p_{transport} = 1 - s/100km$ where s is the distance to the shore in km). This approach is similar to Jambeck et al. [4], who included the mismanaged waste produced by the population within 50 km from the shore. They then applied conversion rates of 15%, 25% and 40% from mismanaged waste to marine litter to estimate the possible range of marine input. In our approach, we calculate the probability to enter the environment by losses to the environment from MSW_T (from both the collected and uncollected waste), as opposed to Jambeck et al. approach to assess conversion rates. We argue that plastic in the environment that is closer to the shore has a higher probability to enter the ocean than waste further inland, while Jambeck et al. assume that all of the mismanaged waste reaches the ocean. We expect our results to be similar to Jambeck et al. because the integrated likelihood corresponds to the transport of 100% of the mismanaged MSW within 50 km from the shore ($\int_0^{100} \frac{x}{100} dx = \int_0^{50} 1 dx$).

Riverine transport
Similarly to the land-based transport, plastic waste in a distance of 20 km is assumed to have a chance to enter the river and this chance is proportional to the distance to the river. Furthermore, waste that enters the river upstream is considered to have a smaller chance to be transported all the way to the ocean than waste that enters further downstream. This probability was set to 100% at the river mouth and sinks gradually to 0% at a distance of 1'000 km upstream ($p_{transport} = 1 - s/1000km$ where s is the distance to the ocean in km). This approach is similar to Jambeck et al. The probability of mismanaged waste of a certain cell to reach the ocean by riverine transport was thus calculated by multiplying the probability to enter the river by the probability to be transported by the river. Thus, mismanaged plastic waste produced near a river in countries that are far from the shore will have a lower chance to end up in the ocean compared to mismanaged plastic waste produced near a river in a country closer to the ocean.
With our river transport model, all types of plastic that enter the river will be transported. In reality, not all plastics will float. Due to their densities, only PE-HD, PE-MD (both 0.94-0.97-0.92 g/ml [74]), PE-LD, PE-LLD (both 0.89-0.93 g/ml

[74]) and PP (0.85-0.92 g/ml [74]) will float in fresh water (1 g/ml) or salt water (1.025 g/ml [2]). PVC (both 1.16-1.41 g/ml [74]) and PET (both 1.38-1.41 g/ml [74]) are too dense. However, also these plastic types could float with a respective shape (e.g. PET bottles). As plastic products are often mixed with fillers which changes their density [2], predictions of the floating behaviour of plastic waste are difficult. However, non floating plastics can still get transported. Hurley [75] found that flooding cleared 70% of all microplastics in river sediments. Thus, we assume that all plastics get transported and do not differentiate between the types in riverine transport. By decreasing the probability of transport with increasing distance from the sea we include losses on the surface (dams, beaching, entrapment [6]) and sediments.

2.6 Estimates of P_O for Future Projections and Policy Chance Scenarios

With the methods described so far in this chapter we can estimate the worldwide P_O on a 0.1x0.1° raster for the years 1990-2015. Besides the evolution of P_O and its geospatial variances, we also want to use the MLM and WFM to predict marine plastic debris inputs for the future and for several scenarios. These include future projections for the years 2030 and 2050 as well as policy change scenarios affecting the losses from uncollected waste and the waste collection system. We created the following scenarios:

- **Future projections for 2030 and 2050**: We created future estimates to assess the change of predicted MSW_C for 2030 and 2050 compared to 2015, assuming business as usual. We used the UN population estimates from 2015 to 2050 [50] to calculate annual growth rates of the population for each country. We then applied these growth rates to the gridded population density data of 2015 [65]. Thus, we assume that the population of each grid cell grows equally and neglect urbanization. The PVs were extrapolated to the years 2030 and 2050 using linear fits of the available data from 1990-2015 (as described in section 2.3.2). Thus, we create the future projections with the MLM while the WFM represents 2015 conditions.
- **100% collection rate for MSW**: We estimate the effectiveness of policy measures that aim to collect all MSW ($MSW_C = MSW_T$). Therefore, we run the MLM with the 2015 data but assume that all of the MSW is collected ($F_{collection}$ = 100%) and no uncollected MSW is generated. Thus, we only change parameters of the WFM.
- **No losses in the waste collection system or no losses of uncollected MSW**: We estimate the effectiveness of policy measures that aim to avoid all losses from the waste collection and treatment system. Again, we run the MLM with all the 2015 data and assume, that all of the MSW_C remains in the collection system and no loss to the environment occurs. Uncollected MSW still leaks to the environment. To compare this with the contrary, we additionally create a (unrealistic) scenario where none of the uncollected MSW enters the environment while losses from the waste collection system occur. Here we assume that all of the informally disposed MSW is discarded in an environmentally safe way. Again, this changes only parameters of the WFM.
- **Decreased transport distances**: We estimate the effects of higher efforts in the management of MSW_C as well as higher rates of collection of waste that leaked to the environment (community cleaning) or removal of plastic waste from rivers. Therefore, we reduce the maximum distance over which plastic waste can reach the ocean by 50% to 50 km from land to the ocean, 10 km from land to rivers and 500 km along rivers to the ocean. This also changes only WFM parameters.

We estimated P_E and P_O using these future projections and policy measure scenarios. The results as well as a comparison to the year 2015 can be found in section 3.4. Next we describe how we predicted the advection of P_O in the ocean using a Lagrangian simulator.

2.7 Simulation of Plastic Particle Distribution by Ocean Currents

In this section we describe how we estimated the distribution of plastic particles entering the ocean on a 0.1x0.1° grid (P_O). This was done by running PARCELS (Probably A Really Computationally Efficient Lagrangian Simulator) [76] using the current data described in section 2.1. The PARCELS project is an open source set of python classes to create particle tracking simulations using hydrodynamic fields. We begin by describing how we modified the raster containing P_O to allow usage in PARCELS. This is followed by the description of the simulation set-up of PARCELS. In the end of this section we present how we analysed the PARCELS output.

2.7.1 Modification of the particle input location

A modification of our P_O data (0.1x0.1° raster) was needed as the current data is available on a 0.25x0.25° resolution (compare table 3). This is a necessary step because the fine 0.1x0.1° raster includes bays that are considered as land in the coarser 0.25x0.25° current data. Particles released at raster cells of the 0.1x0.1° raster might not be inside a cell of the 0.25x0.25° current data. PARCELS would thus not advect these particles as they are considered to be on land. As the current data is also known to be inaccurate close to the shore [62] we decided to release particles from cells that are adjacent to shoreline cells (e.g. further into the ocean). We thus created a 1x1° raster that contains the P_O of all 0.1x0.1° raster cells. In this new raster, the plastic input of each cell is moved to the closest adjacent cell further away from the shoreline. We obtained the coordinates of each cell that represents at least 100 t of plastic waste from P_O. These coordinates were then used in simulation. Next we describe how the simulation was set-up.

2.7.2 Simulation set-up

We run the PARCELS simulation from 1993 until 2015. We therefore use the current data (only from 1993-2017) as described in table 3) and the coordinates of cells where at least 100 t of plastic enter the ocean (previous section). We add particles every 10 days according to the mass of plastic entering the ocean (100 t = 1 particle, 200 t = 2 particles, ...). Each particle is then advected every 10 minutes for one day using 4th order Runge-Kutta and the new position is determined. A Brownian motion kernel is included to simulate turbulent transport as well as micro scale advection not represented in the 0.25x0.25° hydrodynamic field. Therefore, each particles position is changed by a random value between 0 and 100 m/day in both longitudinal and latitudinal direction. This also ensures that particles with the same start point would have slightly different paths to represent both sub-grid scale motions as well as differences in particle size, shape and buoyancy. At the end of each month, the simulated particle density in the 0.25x0.25° ocean grid is stored. Furthermore, every particle is removed from the simulation after three years. This sink represents the settling of fragments as described by Koelmans et al. [42]. Because the model starts with an uncontaminated ocean in 1993 and it takes three years to reach an equilibrium (when the initial particles get deleted), the PARCELS simulation is able to produce meaningful results for the time period 1996-2015.
In the next section we describe how we interpreted the results from the PARCELS simulation using the set-up described above.

2.7.3 Analysis of particle densities

As mentioned above, a particle in the PARCELS simulation represent 100 tons of plastic. This particle size was chosen to reduce the runtime of the simulation. Because we included a Brownian motion kernel, we assume that an increased number of particles (e.g. 1 particle representing 10 t of plastic) would create smoother density distributions. Therefore, we smoothed the densities obtained from PARCELS. We used the *kde2d* function from the *MASS* package [77] to derive kernel density estimates (KDE). KDEs assume that the probability of each particles location can be described as normally distributed, which represents the uncertainty of the Brownian motion. The sum of all particle probabilities represents the cumulative density function. Thus, the product of the KDE and mass of plastic represented by each particle (100 t) equals the amount of plastic in the ocean. We obtained the standard deviation of each particles normally distributed location using the *bandwidth.nrd* function from the *MASS* package [77]. This function uses the *Gaussian approximation* to obtain a optimal smoothness of the prediction compared to the data. The obtained KDE was then multiplied by the amount of plastic that each particle represents to get an estimate of the plastic density in the ocean.

In the following chapter we will present the results of the analyses that were described in this chapter.

3. Results

In this chapter we present the results of our MLM and the estimates for P_E and P_O for the time period 1990-2015 as well as for our scenarios. We start with presenting the selected NNET and RF model in section 3.1. This is followed by estimates of MSW_C for each country in the years 1990-2015 derived from these MLM (section 3.2). We then show the estimated P_E and P_O (including riverine and coastal inputs) for this same time period in the sections 3.3.1 and 3.3.2 respectively. This is followed by the comparison of P_E and P_O of 2015 to the future projections and policy change scenarios (section 3.4). We then present the estimated densities of marine plastic from the PARCELS particle advection model 3.5. In the end of this chapter, we present the results of the machine learning model stability tests in section 3.6.

3.1 Performance of the Machine Learning Models

Here we present the MLM that we selected as described in section 2.4.4). We first present the NNET model,then the RF model. This is followed by a comparison of both models.

NNET

The error of prediction of the unseen test data of the selected NNET model is presented per country in Figure 10 A. For two out of the eleven countries with the highest plastic input, the error of the prediction is <= 10% (compare table 6). From the four countries with the highest plastic input, only the Philippines (94%) have prediction errors below 100%. While the errors of China (119%) and Sri Lanka (129%) are slightly above 100%, the model predicted 6.5 times higher MSW_C for Indonesia (654%) than observed. Additionally, none of the 10'000 NNET models was able to predict China an Indonesia with an error of less than 94% and 211% respectively. In total, 24 countries have errors >100% and an additional eleven have errors >50% (Fig. 10 A). 19 countries have an error of less than 10%. The NNET model has an average MSE of 0.04 t/person/year and an R^2 of 0.88 (table 6 and Figure 10 B). The MSE of the unseen test data in the CV is 0.045 (Fig. 10 C). No country in South America can be predicted with an error of less than 23% (Ecuador) and no low-income country has a prediction error of less than 317% (Senegal) (Fig. 10 D). Additionally, the prediction error of a country usually decreases with increasing income.

RF

The error of prediction of the selected RF model is presented per country in Figure 11 A. For six out of the ten countries with the highest plastic input, the prediction error is <= 10% (compare table 6). From the four countries with the highest plastic input (compare table 6), China (49%) and the Philippines (25%) have prediction errors below 50% and Indonesia (215%) and Sri Lanka (294%) have errors between 200-300%. However, none of the 10'000 RF models was able to predict these two countries with an error less than 180% and 241% respectively. In total nine out of 125 countries have errors

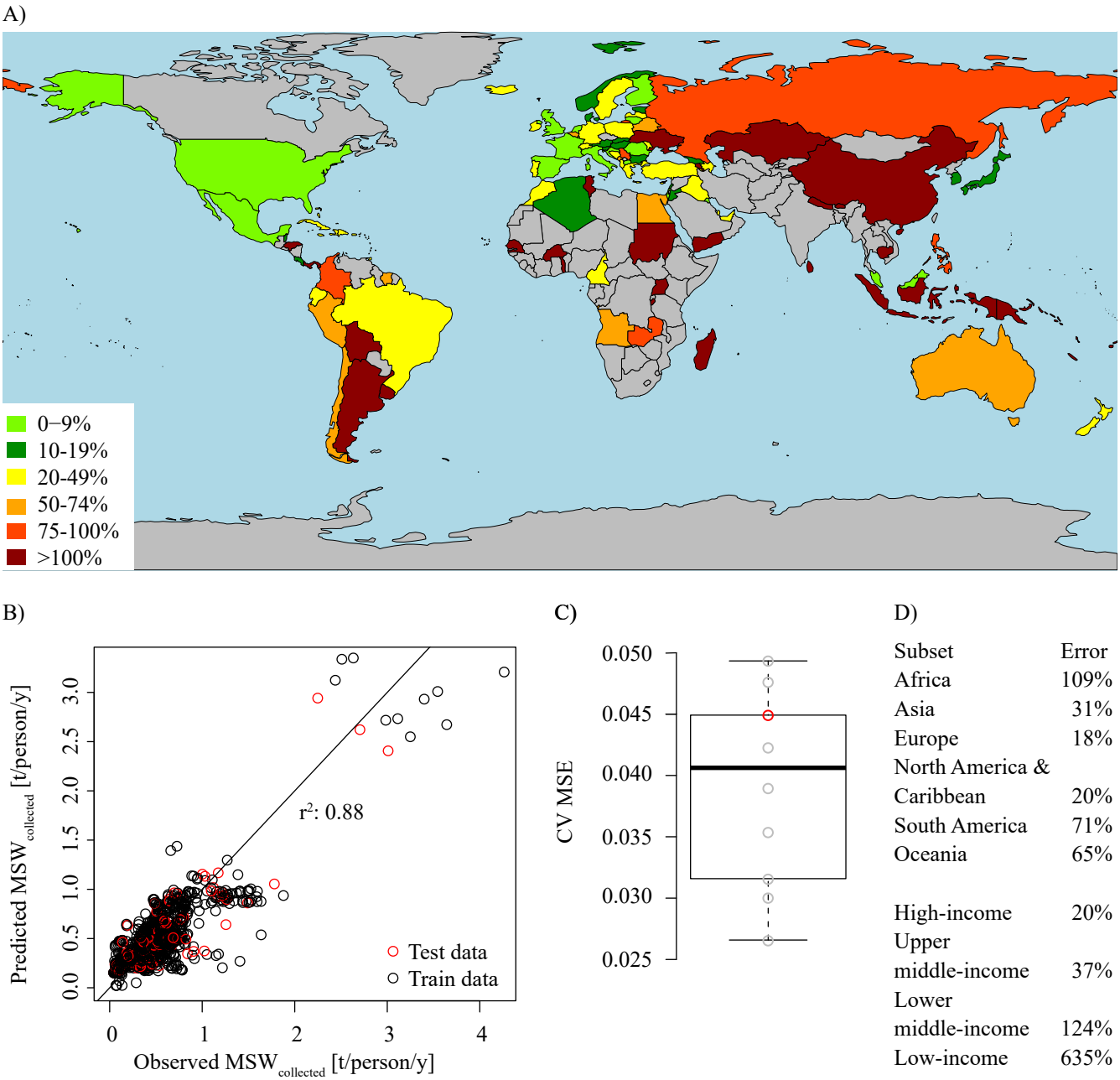

Figure 10. Summary of the selected NNET model. A: Map including the mean error of the predicted versus observed MSW_C in percent Grey countries had no data available. B: Comparison of the predicted versus observed MSW_C for all available data. Red circles represent the test data which was not used in training the RF. C: Boxplot showing the MSE of the 10-fold CV that was performed after training the model. The grey circles represent training data and the red circles represent test data. D: Median error per region and income.

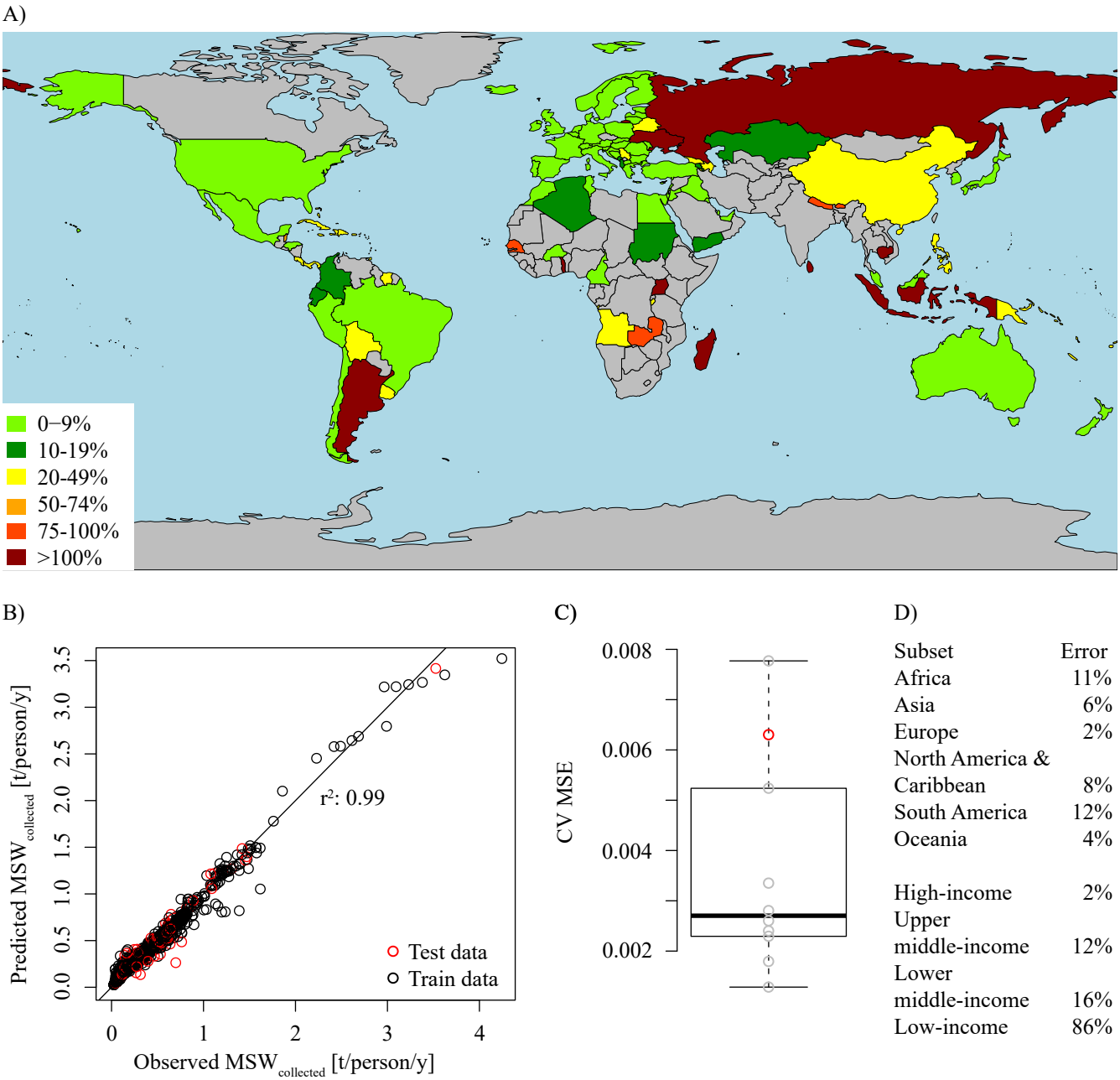

Figure 11. Summary of the selected RF model. A: Map including the mean error of the predicted versus observed MSW_C in percent. Grey countries had no data available. B: Comparison of the predicted versus observed MSW_C for all available data. Red circles represent the test data which was not used in training the RF. C: Boxplot showing the MSE of the 10-fold CV that was performed after training the model. The grey circles represent training data and the red circles represent test data. D: Median error per region and income.

Table 6. Error of prediction versus observation of the selected NNET and RF models for the countries that contribute the most to marine plastic pollution according to Jambeck et al. [4] as well as the MSE and r^2. The error of prediction of these eleven countries with a high impact on P_O was used to select the NNET and RF models that have the best overall performance (section 2.4.4). The column *range* indicates the minimum and maximum error of all 10'000 generated models for that country. An estimate of P_O for our top 10 countries can be found in table 10.

	NNET error [%]		RF error [%]	
	Used	Range	Used	Range
China	119	94–-352	49	41–116
Indonesia	654	211–1'340	216	180–422
Philippines	94	0–269	25	10–103
Sri Lanka	129	0–2'116	294	241–861
Egypt	-63	-19–-85	-4	0–-18
Malaysia	-5	0–-41	-2	0–-8
Algeria	14	0–54	11	0–23
Turkey	-32	-5–-57	-8	-6–-14
Brazil	44	-2–98	10	6–17
Morocco	28	0–83	1	0–55
USA	-2	0.0–-35	0	0–-1
MSE	0.04	0.03–0.08	0.004	0.003–0.009
r^2	0.88	0.92–0.80	0.99	0.99–0.98

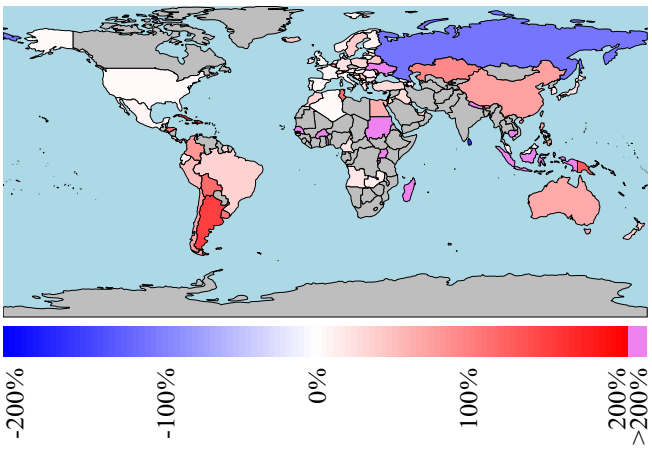

Figure 12. Difference of the NNET and RF model prediction error for each country. Negative values indicate countries that are better predicted by the NNET model, positive values indicate countries that are better predicted by the RF model. Grey indicates countries where no MSW_C data was available.

>100% and an additional four have errors >50%. 39 countries have an error of less than 10%. The RF model has an average MSE of 0.0036 and an r^2 of 0.99 (compare table 6 and Figure 11 B). The MSE of the unseen test data in the CV is 0.0063 t/person/year (Fig. 11 C). For each continent there are several countries with an prediction error <10% (Fig. 11 D). Also, for each income class, there is at least one country with a prediction error of less than 10%. However, the prediction error of a country usually decreases with increasing income.

Comparison of the NNET and RF model

After describing the selected NNET and RF model we now compare them. Figure 12 shows the difference of the NNET and RF prediction error. The RF model is more successful in predicting the MSW_C than the NNET model in 100 out of 106 countries where MSW_C data was available. The six countries where the NNET model has the smaller error in prediction include Andorra, Belize, Costa Rica, Georgia, Russia and Sri Lanka. However, for Andorra, Costa Rica and Georgia, the difference in the NNET and RF prediction error is below 6%. The error is bigger for Belize, Russia and Sri Lanka, where the NNET model has an error of 15%, 76% and 129% respectively while the RF model has an error of 38%, 182% and 293% respectively. Thus, the RF model prediction is always significantly more accurate than the NNET, except in 3 out of 106 countries where the NNET error is significantly lower. We therefore dropped the NNET model for the final analysis and only used the RF model to calculate MSW_C and thereof P_E and P_O. In the next section we will present the estimates of MSW_C derived from the RF model.

3.2 Predicted MSW_C

Here we present our estimates for MSW_C which we derived using the RF model presented in the previous section. We dropped the NNET model because the RF model was superior in the prediction performance (Fig. 12).

The RF models predicts an increase in the average worldwide MSW_C from 0.28 t/person/year in 1990 to 0.35 t/person/year in 2015 (table 7). Of all countries, Kuwait is predicted to collect the highest amount of MSW in 2015 (3.29 t/person) and Burundi collected the least MSW with 6 kg/person in 2015. The 20 countries with the highest predicted MSW_C in 2015 are located on all continents except Africa and South America (Fig. 13 A). All of these are high-income countries, where the median error of prediction is 2% (Fig. 11 D). Only two of the 20 countries with the lowest predicted MSW_C are not located in Sub Saharan Africa (Laos and Afghanistan) and 14 of them are low-income countries.

The USA collect the most MSW (0.65 t/person) and Indonesia the least (0.10 t/person) from the 20 countries that have the highest impact on marine plastic debris according to Jambeck et al. [4]. China, the country with the highest impact on marine plastic debris [4], has a predicted MSW_C of 0.22 t/person and 0.26 t/person in 1990 and 2015 respectively.

Figure 13 A shows the averaged predicted values of MSW_C

from the RF model for the years 1990-2015. In all countries in Western Europe (except Portugal), North America and Oceania, the amount of MSW_C has remained constant or decreased slightly. At least 0.4 t of MSW_C per capita and year are predicted in these countries. The average MSW_C in these continents has remained almost constant during this time period (compare table 7 and Fig. 13 B). In Africa, no country collected more than 0.4 t MSW per person in 1990 and most countries are predicted to have a collection of no more than 0.1 t per person and year. The continental average is thus by far the lowest at 0.09 t/person (table 7). However, several countries in Sub-Saharan Africa are predicted to have increases of MSW_C in the time period 1990-2015 (Fig. 13 B). This is also apparent in the 2015 continental average which has increased to 0.15 t/person (table 7). A similar trend can be seen in South America, where average MSW_C has risen from 0.23 t/person in 1990 to 0.32 t/person in 2015 and all countries except Colombia collect the same amount or more waste in 2015 compared to 1990 (Fig. 13 B). Asia has the strongest increase in the average MSW_C, from 0.29 t/person in 1990 to 0.47 t/person in 2015 (table 7). However, while many countries in South Asia, including the Arabian Peninsula, have rising MSW_C rates, Russia and Laos (-0.20 kg/person/year and -0.22 kg/person/year respectively) are predicted to have a declining rate (Fig. 13 B). However, the prediction error of Russia is 182% (Fig. 11 A) and this result has to be considered with caution.

From all 174 countries, 12 exhibit a decrease in MSW_C of at least 0.1 t/person/year between 1990 and 2015 (Fig. 13 B). Eight of these are high-income countries, three and one are upper and lower middle-income countries respectively. In total, 53 countries show a decrease (Fig. 13 B). Thereof, 38 show a decrease of <1%. Monaco exhibits the largest decrease by 43% from 1.32 t/person/y in 1990 to 0.89 t/person/y in 2015.

Table 7. Summary of the modelled MSW_C in 1990 and 2015 for the three countries with the highest collection as well as averaged for the income classes, the continents and worldwide.

	Subset	MSW_C [t/person/year]	
		1990	2015
	World	0.28	0.35
Top 3	Kuwait	2.13	3.29
	Qatar	0.65	1.50
	Bermuda	1.26	1.22
Income	High-income	0.54	0.60
	Upper middle-income	0.24	0.32
	Lower middle-income	0.12	0.23
	Low-income	0.05	0.09
Continents	Africa	0.09	0.15
	Asia	0.29	0.47
	Europe	0.47	0.47
	North America & Caribbean	0.38	0.37
	South America	0.23	0.32
	Oceania	0.52	0.49

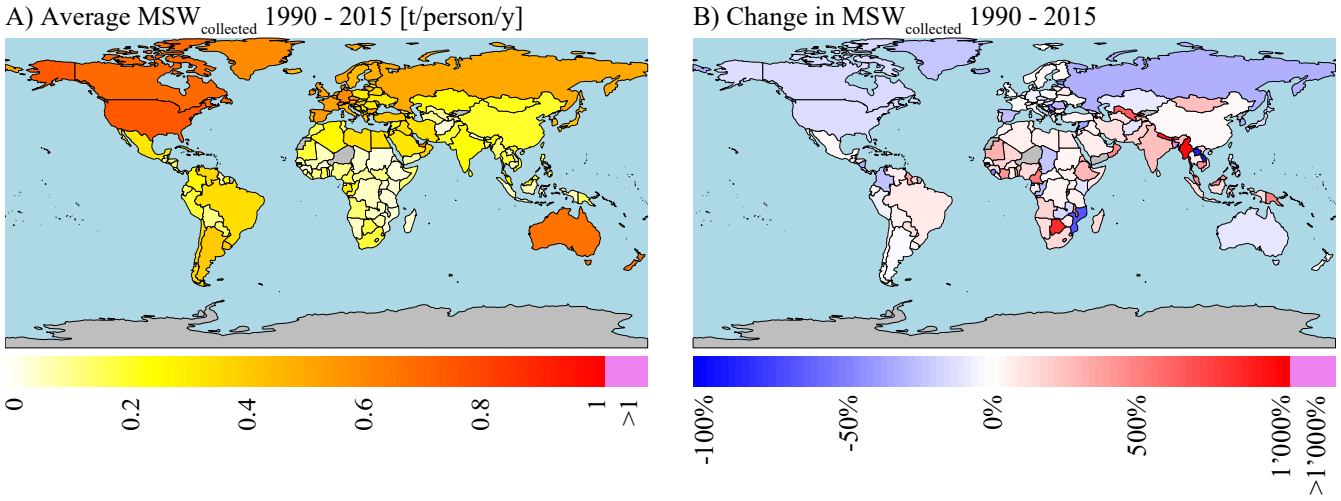

Figure 13. Predicted MSW_C from the RF model averaged from 1990 - 2015 for each country and the relative change in 2015 compared to 1990. Please note the tenfold difference in the scale of the change plots. Countries with a MSW_C >1 t/person/year (Singapore, Qatar, Monaco, Bermuda and Kuwait) or a relative change >1'000% (Lesotho and Bangladesh) are marked in violet to keep the colour gradients in scale.

3.3 Predicted Plastic Fluxes to the Environment and the Oceans

In this section we present our predictions for P_E and P_O, derived from the WFM presented in section 2.5 and Figure 8. We start with the estimates of P_E and then present the estimates of P_O broken down to riverine and coastal inputs.

3.3.1 Predicted P_E

The WFM suggests that the amount of worldwide P_E has continuously risen from 31.8 Mt in 1990 to 51.0 Mt in 2015 (table 8). While the plastic that entered the environment in 1990 originated in equal amounts from collected and uncollected MSW (15.9 Mt each), this fraction has changed over time. In 2015, 58% originated from MSW that entered the collection system while 42% originated from MSW that was not collected.

3.3.2 Predicted Plastic Fluxes to the Oceans

Here we present the estimates for P_O separately for riverine and coastal inputs and as a total.

Fluxes from rivers

The WFM (compare Fig. 8) predicts that 0.47 Mt (3%) of the plastic that enters the ocean in 1990 (13.4 Mt) originated from riverine sources (table 8). The fraction of riverine sources is predicted to double (6%) until 2015. Figure 14 A shows the evolution of riverine plastic flux to the ocean from 1990 to 2015. In 1990, rivers in Asia and Europe are predicted to have released 225 kt (48%) and 162 kt (34%) respectively into the oceans. All the other rivers accounted for 95 kt. In 2015, the highest fluxes are predicted in Asia with a total of 739 kt (70%), followed by Africa and Europe with 135 kt (13%) and 114 kt (11%) respectively and 73 kt are released in the remaining continents combined. Europe is the only continent where the riverine input decreased in the study period (by 51%). However, the input in Asia and South America tripled while input in Africa and North America doubles during the same time period. Table 9 shows the input of the top 10 ranked rivers in 2015. Eight of these ten rivers are located in Asia and five flow through China (Yellow River, Yangtze, Xi River, Red River and Mekong). Figure 15 shows the global riverine input in 2015. It can be seen that the rivers with the highest plastic loads ($>$10 kt/year) are located in Asia, Africa and Europe while no rivers in the Americas or Oceania has loads $>$10 kt/year. Furthermore, a North-South gradient can be seen. Rivers that flow to the Arctic Sea carry $<$100 t/year. The river Ob in Russia is the only exception with $>$100 t/year. Also rivers in Australia and New Zealand as well as Alaska and Canada usually carry $<$100 t/year. All other estimated river loads are $>$100 t/year. In 1990, 61% (295 kt) of all riverine plastic waste originated from lower middle-income countries. This fraction has remained constant until 2015 (60%, 638 kt). However, while the input from high-income and low-income countries has only risen from 73 kt to 95 kt and from 22 kt to 35 kt respectively, the input from upper middle-income countries has tippled from 92 kt to 292 kt. Together, upper and lower middle-income countries account for 88% of the global riverine marine plastic input in 2015.

Fluxes from the coast

The input from coastal sources (compare Fig.14 B) is predicted to increase from 13.0 Mt to 17.7 Mt while accounting for 97% and 94% respectively of P_O in the years 1990 and 2015 (table 8). Predictions for countries in Europe, North America an Oceania remained relatively constant over time, but many African, Asian and South American countries have increased their plastic input (Fig. 13 B). African countries emitted 279 kt (2% of global coastal input) in 1990 and 728 kt (4% of global coastal input) in 2015 and thus doubled their fraction of global coastal inputs (table A.3). The total coastal plastic input from South American countries increased from 213 kt (2%) to 592 kt (3%), North America from 331 kt (3%) to 587 kt (3%) while Europe decreased from 900 kt (7%) to 732 kt (4%). The coastal plastic input from Asian countries rose from 11'217 kt (86%) to 15'015 kt (85%). China releases 8'539 kt and is thus accountable for more than half of the Asian coastal plastic input and 48% of the global coastal input in 2015.

All low-income countries (representing 11% of the global population in 2015) combined account for 1% of the coastal plastic waste flux to the ocean in 2015 and high-income countries (31% of the global population) account for 9% (table A.3). Thus, upper and lower middle-income countries account for 55% and 35% respectively (while representing 14% and 44% of the global population respectively). However, the total fluxes from low-income and lower middle-income countries have tripled since 1990 while the upper middle-income countries have remained relatively constant (9'319 kt in 1990 and 9'762 kt in 2015).

Total plastic influx to the oceans

We predict an increase in total marine plastic input from 13.4 Mt in 1990 to 18.7 Mt in 2015 (table 8). Table 10 shows the ten countries that we estimate to produce the most marine plastic litter in 2015. This tables also shows the evolution of the marine plastic litter input of these countries. China alone produces 8'770 Mt of marine plastic debris in 2015, which accounts for 47% of the worldwide input in 2015. Compared to the country with the second highest input (Indonesia), China accounts for six times more marine plastic litter. From the top ten countries, only Brazil and the USA are not located in Asia and out of the Asian countries, only Bangladesh and India are not in South East Asia. Combined, the top 10 emitters account for 11.1 Mt in 1990 and 15.1 Mt in 2015 which represents 83% and 81% of the worldwide input. From 1990 to 2015, only the input of Europe decreased from 1'063 kt to 847 kt (table A.4). In Africa (1990: 240 kt, 2015: 864 kt) and North America (1990: 359 kt, 2015: 637 kt) the input roughly doubled while South America (1990: 220 kt, 2015: 615 kt) tripled its plastic flux. The input of Oceania increased from 31 kt to 37 kt, which represents 0.2% of the worlds input.

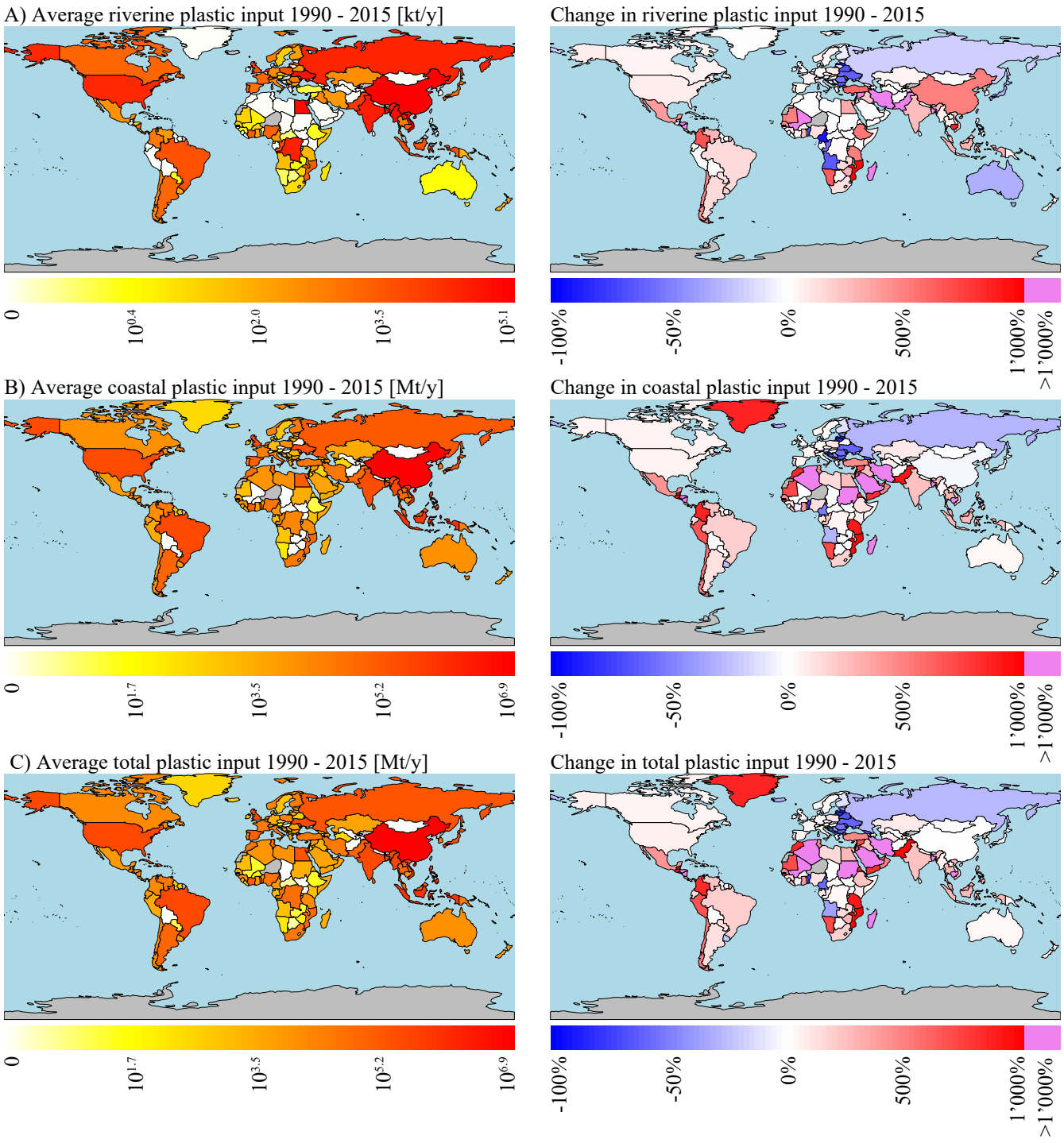

Figure 14. Predicted marine plastic input averaged for the years 1990 - 2015 and the relative change in 2015 compared to 1990. A) Modelled input from riverine sources per country. B) Modelled input from coastal sources per country. C) Modelled input from both coastal and riverine sources per country. Please note the log-scale in the average input and the tenfold difference in the scale of the change plots. Countries with a relative change >1'000% are marked in violet to keep the colour gradient in scale.

Table 8. Summary of the modelled fluxes of plastic to the environment and the ocean for the years 1990, 1995, 2000, 2010 and 2015. We separate the P_E by inputs from the waste collection system and from uncollected waste. The P_O is separated by inputs from coastal areas and from riverine sources.

	1990	1995	2000	2005	2010	2015
Flux from MSW_C to P_E [Mt]	15.9	17.5	20.0	21.6	27.1	29.5
Flux from uncollected MSW to P_E [Mt]	15.9	17.3	19.4	19.8	22.3	21.5
Total P_E [Mt]	31.8	34.9	39.5	41.4	49.4	51.0
Flux from P_E to P_O by rivers [Mt]	0.47	0.55	0.72	0.81	1.07	1.1
Flux from P_E to P_O by coastal areas [Mt]	13.0	13.8	15.7	15.1	16.5	17.7
Total P_O [Mt]	13.4	14.4	16.4	15.9	17.5	18.7

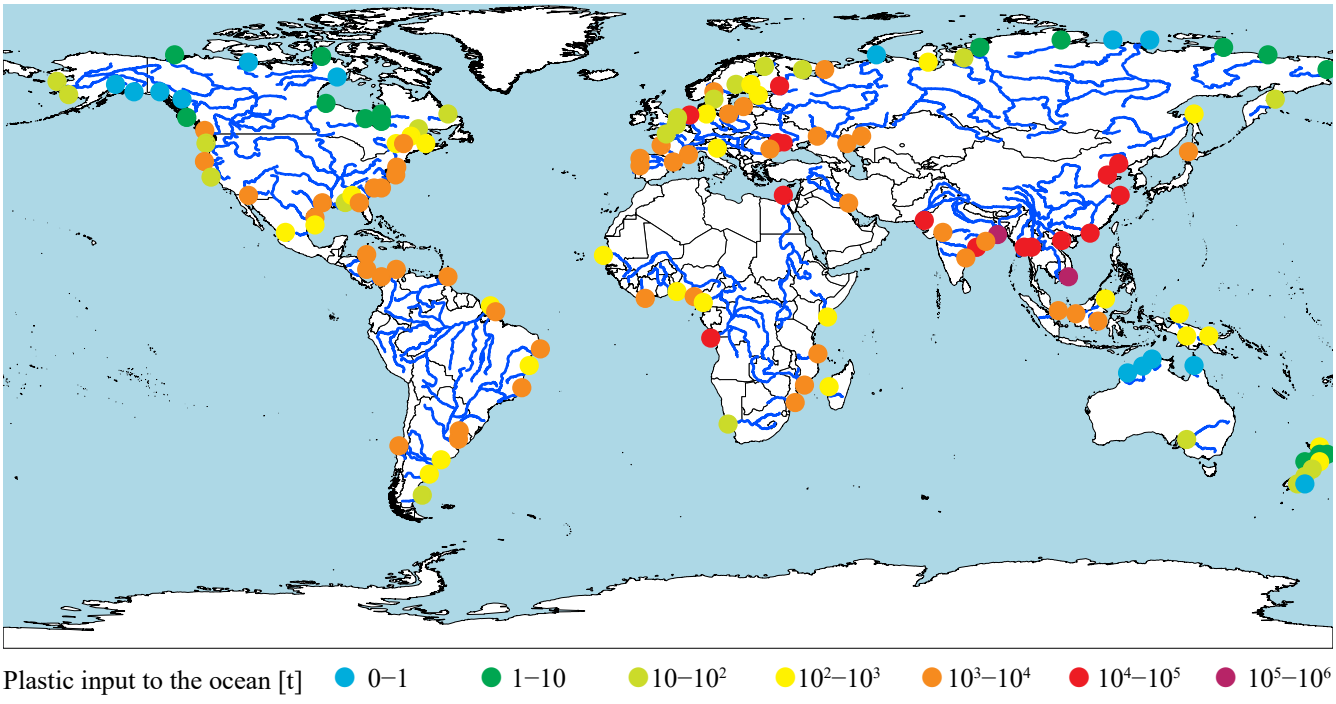

Figure 15. Predicted riverine plastic input to the oceans in 2015. In total, 342 rivers transported 1.1 Mt of plastic to the ocean (compare table 8). Please note the log-scale. A more detailed description can be found in section 2.5.3

Table 9. List of the 10 rivers with the largest estimated plastic input to the ocean in kt in 2015, ranked by their 2015 plastic input. The number in the column *world* represent the fraction of the total riverine plastic input in 2010.

Name	Input 2010 [kt]	World
Ganges-Brahmaputra	145'691	13.5%
Yellow River (Huang He)	104'282	9.7%
Nile	99'924	9.3%
Yangtze (Yangzi Jiang)	89'358	8.3%
Indus	77'087	7.2%
Pearl River (Zhujiang)	62'684	5.8%
Red River (Yuang Jiang)	60'430	5.6%
Irrawaddy	42'556	3.9%
Mekong	34'997	3.2%
Congo	30'202	2.8%

As with the coastal plastic inputs, all low-income countries combined account for 1% of the total plastic waste flux to the ocean in 2015 and high-income countries account for 9% (table A.4). Thus, upper and lower middle-income countries account for 53% and 37% respectively. However, the total fluxes from low-income and lower middle-income countries have tripled since 1990 while the upper middle-income countries have remained relatively constant (9'410 kt in 1990 and 10'055 kt in 2015).

After describing the results of our estimates of P_O for the time period 1990-2015, we present the estimates of P_O we made for the future projections and the policy change scenarios in the next section.

3.4 Scenarios

Besides the P_O estimates for the years 1990-2015, we also used the MLM to predict future projections and the WFM to predict policy change scenarios. This includes projections for the years 2030 and 2050 as well as scenarios including a 100% collection rate of MSW, no losses of plastic to the environment and a decreased transport distance of plastic litter (as described in section 2.6). Here we first present the estimates of the future projections, followed by the predictions for the policy change scenarios.

3.4.1 Future projections

Here we present the projections for the years 2030 and 2050 (assuming business as usual) and compare them to the year 2015. These were created by extrapolation the population density and the PVs used in the RF model presented in section3.1. The PVs and the population density were extrapolated as described in section 2.6.

Table 10. Summary of the modelled P_O from 1990 to 2015 for the ten countries with the highest plastic input to the oceans in 2015, ranked by the 2015 values. A detailed description can be found in section 2.5.

Country	Total marine plastic input [kt]					
	1990	1995	2000	2005	2010	2015
China	8'854	8'888	9'512	8'067	8'632	8'770
Indonesia	395	634	605	427	783	1'431
Vietnam	369	473	704	1'115	1'220	1'197
Timor-Leste	529	648	800	896	965	1'096
Philippines	231	358	609	785	769	782
India	122	220	258	322	293	404
Brazil	135	131	250	433	402	377
Myanmar	270	293	335	162	291	372
Bangladesh	16	48	191	223	312	357
USA	232	245	287	300	312	328

Projection for 2030

For 2030, the RF model predicts a P_O of 19.7 Mt, which is a 5% increase compared to 2015 (table 11). This increase mostly originates from Africa, which increases its input by 50% to 1'210 kt (+346 kt) and Asia, whose input increases by 1.7% (+282 kt). Nine of the ten countries with the highest input are still the same as in 2015 (compare table 10). The input of the USA decreases from 238 kt in 2015 to 203 kt in 2030. Meanwhile, Pakistan increases its input from 149 kt to 369 kt and thus replaces the USA in the list of the biggest emitters in 2030. Also China decreases its input by 51% (8'770 kt in 2015) to 5'816 kt. Nevertheless, China is till on the higher rank than Indonesia, whose input increased by 39% to 1'995 kt. We project a 41% increase in riverine input (from 739 kt in 2015 to 1043 kt in 2030) in Asia while the fluxes in South America decrease (from 21.9 kt in 2015 to 9.6 kt in 2030). Riverine transport from lower middle-income countries could double (to 1'169 kt) compared to 2015 while the input from upper middle-income countries might increase 1.5 times (to 173 kt). The flux from the shore is predicted to remain constant for Asia (15'015 kt in 2015 ad 14'993 kt in 2030), while it increases by 39% (from 723 kt to 1'013 kt) in Africa and by 23% (733 kt to 900 kt) in Europe. The other continents change by $<$10%.

Projection for 2050

For 2050, we predict a P_O of 21.6 Mt, which is a 16% increase compared to 2015 (table 11). This increase mostly originates from Asia, which increases its input by 12% (+1'868 t) and Africa, whose input increases by 60% to 1'382 kt (+ 518 kt).The same countries are atop of the top 10 list as in 2030. While China is still leading with 6'403 kt (decrease of 27% compared to 2015), the Philippines are second with 2'137 kt

(increased by 173% compared to 2015). The riverine fluxes increase in Africa (from 135 kt in 2015 to 209 kt in 2050) and Asia (from 739 kt in 2015 to 1'279 kt in 2030) the input of the other continents changes <10%. The plastic input from the shore increases in Africa (+445 kt to 1174 kt), Asia (+1'337 kt to 16'342 kt) and North America (+155 kt to 578 kt). We project that the rise from 2015 to 2050 in the total worldwide marine plastic input (2.9 Mt) mostly originates from Asia (+1'868 kt) and Africa (+518 kt) the other continents only show moderate increases. In the income classes, a shift from the upper middle-income countries (-2'764 kt) towards lower middle-income countries (+4'961 kt) might arise.

Comparison of the 2030 and 2050 projections

We project an increase of P_O for the years 2030 and 2050 (table 11). The increase in the period 2030-2050 (+1.7 Mt total, +85 kt/year) is larger than in the period 2015-2030 (+1 Mt total, +67 kt/year). Figure 16 shows that we project a strong increase in P_O originating from lower middle-income countries between 2015 and 2030. However, in the same time period we also predict a decrease in P_O originating from upper middle-income countries. Thus, the total P_O increase is subdued. In the time period 2030-2050, we project increasing P_O for both lower and upper middle-income countries. Thus, the global P_O shows a larger increase in this period. Figure 16 also shows that the P_O inputs of the low-income and high-income countries only make up a small fraction of the global input.

3.4.2 Policy change scenarios

Here we present the estimates for the scenarios. We created three scenarios, simulating policy measures that aim to collect all waste, decrease losses to the environment and reduce the transport to the ocean. These scenarios were created by changing parameter values of the WFM presented in section2.5. We used the MSW_C predictions of 2015 and thus compare to this year. A detailed description can be found in section 2.6.

No uncollected MSW

In this scenario we assume that 100% of the MSW is collected ($MSW_C = MSW_T$). Table 11 shows, that we predict a reduction of P_E by 15% to a total of 43.4 Mt (as compared to the regular value of 51.0 Mt in 2015). We predict that P_O is reduced by 10% to 16.9 Mt (18.7 Mt in 2015). A decrease can be seen in both riverine transport (-0.3 Mt) and influx from the shore (-1.7 Mt). Thus, we predict that collecting all the MSW produced in 2015 would only decrease the P_E by 15% and the P_O by 10%. Comparing the continents reveals that the biggest change in P_O is predicted in Asia (-1.5 Mt) while the other continent show reductions of less than 0.1 Mt. Considering the income classes, both upper and lower middle-income countries could reduce their predicted marine input by 0.8 Mt while high and low-income countries have lower decreases. Thus, increasing the $F_{collection}$ might only have a significant result on the amount of P_O in Asian middle-income countries.

No losses

In this scenario we assume that no flux to P_E occurs from either the waste collection system or from uncollected waste. In the case that no losses occur from the waste collection system, we predict a P_E of 21.5 Mt and a P_O of 7.7 Mt (table 11). This represents a reduction of 58% and 59% respectively compared to the regular 2015 values. In the case where no losses occur from uncollected waste, we predict a P_E of 29.5 Mt and a P_O of 11.1 Mt. This represents a reduction of 42% and 41% respectively compared to the regular 2015 values. Furthermore, our predictions indicate that 40% (0.4 Mt) of the riverine plastic waste originates from MSW_C while 60% (0.6 Mt) originate from uncollected waste. From the coastal fraction of P_O, 41% (7.2 Mt) originate from MSW_C while 59% (10.5 Mt) originate from uncollected waste. Again, the biggest effect of this measure can be seen in Asia, which contributes to 90% (6.9 Mt) of the reduced plastic inputs. Also reducing the losses in the middle-income countries would result in an 90% reduction. Worldwide, minimizing the losses from the waste collection system could reduce the worldwide P_E and P_O by up to 60%.

Halved transport distances

In this scenario we assume that the probability of transport of plastic is halved (e.g. due to collection of P_E). Table 11 shows that this could lead to a reduction of P_O by 29% to 13.3 Mt (compared to the regular value of 18.7 Mt in 2015). Efforts in removing plastic from the environment could reduce the riverine plastic input by 55% (to 0.5 Mt), and the coastal input by 28% (to 12.8 Mt). This policy measure shows no major differences among the continents for riverine and coastal inputs.

3.5 Predicted Concentrations of Floating Marine Plastic

Here we present the estimated changes for the concentrations of marine plastic debris in the years 1996 and 2015. Figure 17 A shows the yearly averaged plastic concentrations in the oceans have changed from 1996 to 2015. Oceans that show no difference include the South Pacific and the Southern Ocean. All other oceans show zones of changed concentrations. In the North Pacific, North and South Atlantic and the Indian Ocean, zones of increase are related to the known aggregation zones of the gyres. These oceans also include minor zones of decrease. The largest increases can be seen in the southern Indian Ocean, in the North Atlantic, in the Arctic Sea near Scandinavia and western Russia and in the Seas around Japan. We predict a major decrease of plastic concentrations in the the North Sea, the Baltic Sea and in the western North Atlantic (Bay of Biscay). Comparing the mean concentration of the years 1996 and 2015 (Fig. 17 B&C) reveals that the shape of the aggregation zones have changed from 1996 to 2015. This results in the small zones of decrease found in most Oceans. The change of the shape of the aggregation zones can be explained by slightly different currents in these years.

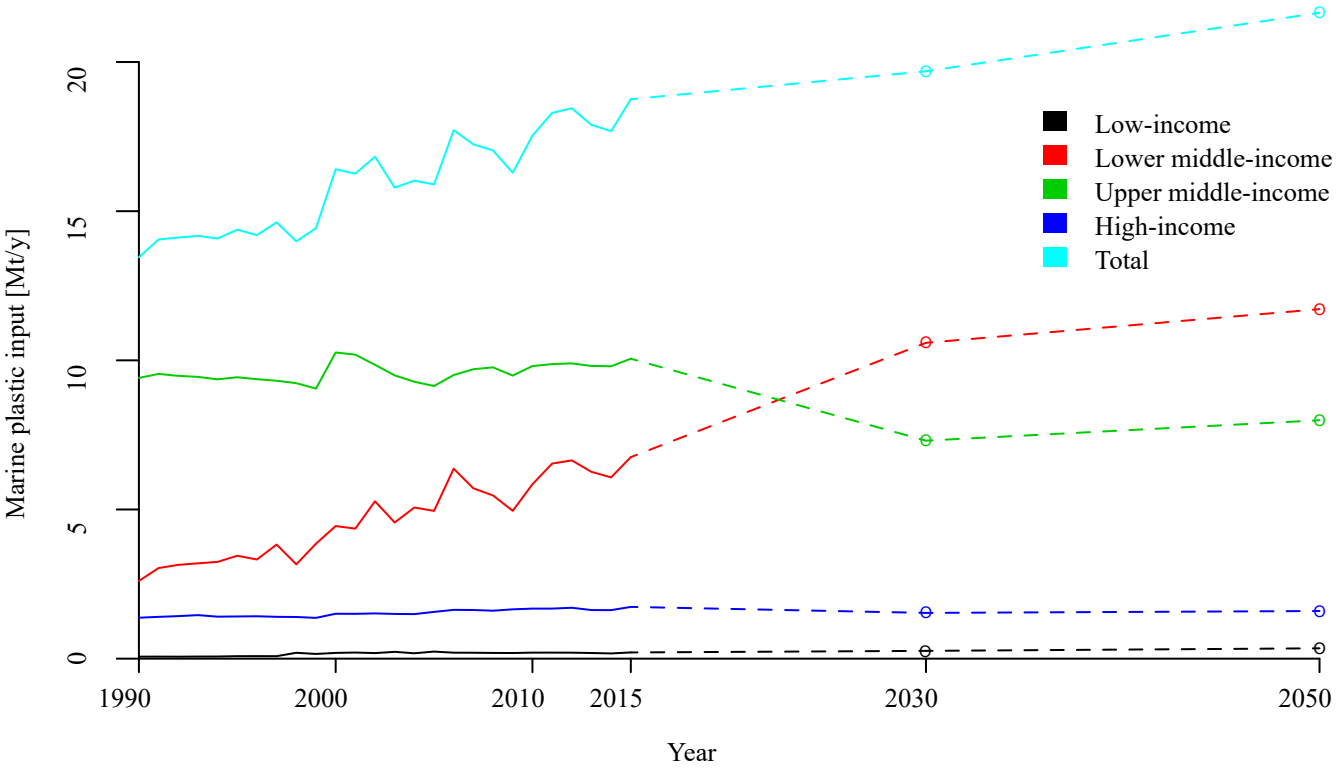

Figure 16. Estimated marine plastic input for the years 1990-2015 (solid lines) including projections for 2030 and 2050 (marked with points). The estimates are provided for the four income classes and for the whole world. The dashed lines represent the trends between 2015 and 2030 and between 2030 and 2050 but are not based on estimates for the years in-between.

Table 11. Summary of the modelled fluxes of P_O and P_O for the future projections of the years 2030 and 2050 as well as for the policy change scenarios. We include the year 2015 as a comparison. The future projections are derived using the MLM with projected PVs and population density and the 2015 WFM. The policy change scenarios are derived using the 2015 MLM estimates and changes in the WFM (see section 2.6).

	2030	**2050**	**All coll.**	**No loss in**		**Distance**	**2015**
				uncoll.	coll.		
Plastic flux from collected MSW to the environment [Mt]	32	36.8	43.4	29.5	0.0	29.5	29.5
Plastic flux from uncollected MSW to the environment [Mt]	22.4	26.5	0.0	0.0	21.5	21.5	21.5
Total plastic input to the environment [Mt]	54.4	63.3	43.3	29.5	21.5	51.0	51.0
Plastic flux from rivers to the ocean [Mt]	1.5	1.7	0.8	0.6	0.4	0.5	1.1
Plastic flux from coastal areas to the ocean [Mt]	18.2	19.9	16.0	10.5	7.2	12.8	17.7
Total plastic input to the ocean [Mt]	19.7	21.6	16.9	11.1	7.7	13.3	18.7

Overall, the only major decrease in plastic concnetrations is predicted for the North Sea and the Baltic Sea. As no current data is available for the Mediterranean, we excluded plastic particles in this Ocean.

3.6 Model Stability

Here we present the results of the stability test performed on the NNET and RF models as described in section 2.4.3. We tested both the bias of the training data size and the bias of in the training data selection.

Both models have a better prediction performance the bigger the training dataset is (Fig. 18 A&B). Generally, the more data is used for training, the better is the model performance. Both MLM perform best when the training data consists of 90% of the available data (white lines). Additionally, the RF is more sensitive to the size of the training dataset than the NNET. However, the effect on the error is relatively low when varying the size between 80% and 50% compared to the models trained on 40% or less.

A strong variation in the error of the NNET and RF model prediction (MSE) is observed when the training data is biased (Fig. 18 C&D). Both MLM perform best when trained on the high-income countries and perform worst when trained on the lower middle and low-income countries. Overall, both models behave similarly when trained on the same data. While the size of the training data is between 1% (Oceania dataset) and 64% (high-income dataset), the MSE is always lager than for randomly sampled training data (Fig. 18 A&B). Thus, MLM trained on biased data will have a larger MSE compared to MLM trained on randomly sampled training data. The effect of the cluster (where C1 represents countries with the lowest GDP and C7 those with the highest GDP) does not follow the trend observed with the income where "rich" countries perform better. When the models are trained on data from Europe, the prediction is worse than when the training is performed on other continents.

Contrary to the above described bias in the training data size, the effect of the size of the training data is of less importance. Training on the 17 data points from Oceania results in an overall better prediction for worldwide data than training on 773 points from the countries in Europe.

In the following chapter we will discuss the results that were described in this chapter.

4. Discussion

In this chapter we will discuss the results that we presented in the previous chapter. We will first compare the results of MSW_C and P_O and assess implications. We then compare our main results with findings from the literature. This includes our estimated MSW_C rates as well as the estimated P_O by riverine sources and in total. Following, we propose methods to assess the error of our estimates. Then we show how the RF model error might affect our estimates and propose suggestions for how this could be changed. In the end we evaluate if the goal we presented in the introduction were reached.

4.1 Implications of the Results

Here we compare the results of MSW_C and P_O as presented in the previous chapter. We show trends and implications and present possible explanations. We suggest possible solutions to tackle the problem of marine plastic debris. We start by describing the regional and temporal changes of P_O. Following we discuss the influence of the income on P_O. Then we compare the effect of increasing $F_{collection}$ and decreasing losses from the waste collection system. In the end of this section we discuss the effect of the transport distance on P_O.

4.1.1 Geographical Differences over Time

Our results show that the predicted P_O has mostly increased from 1990 to 2015 in African, South Asian and South and Central American countries (Fig. 14). Eastern European countries show the largest decrease. These results align with the estimates of MSW_C for 1990 and 2015 (table 7), where we predict increases in Africa, Asia and South America. For Europe, North America, and Oceania we predicted constant values. The decrease in Eastern Europe can be explained by the $F_{collection}$ rates which have increased from 63% in 1990 to 82% in 2015. Thus, the MSW_T has decreased in Eastern Europe (by 11%) and the decreased losses from uncollected waste have lead to lower P_O.

4.1.2 Influence of the Income on P_O

We predict that in 2015, the middle income countries (58% of the global population) account for 90% of the global plastic input to the oceans (compare Fig. 16). This is in agreement with Jambeck et al. [4] who also identify 16 of the 20 countries with the highest plastic inputs to the ocean as middle-income [4]. This could be explained by fast economic growth in these countries while the waste management system lags behind the increasing MSW production [37]. When comparing our predictions for 2030 and 2050, the input of the current upper middle-income countries could reach their "peak waste" in this time period (compare Fig. 16). On the other hand, the input from the current lower middle and low-income countries will continuously increase. Assuming further economic growth, the countries could rise in their income classification and follow the same trend as today's high-income countries. This implies that P_O will further increase in the next 30 years

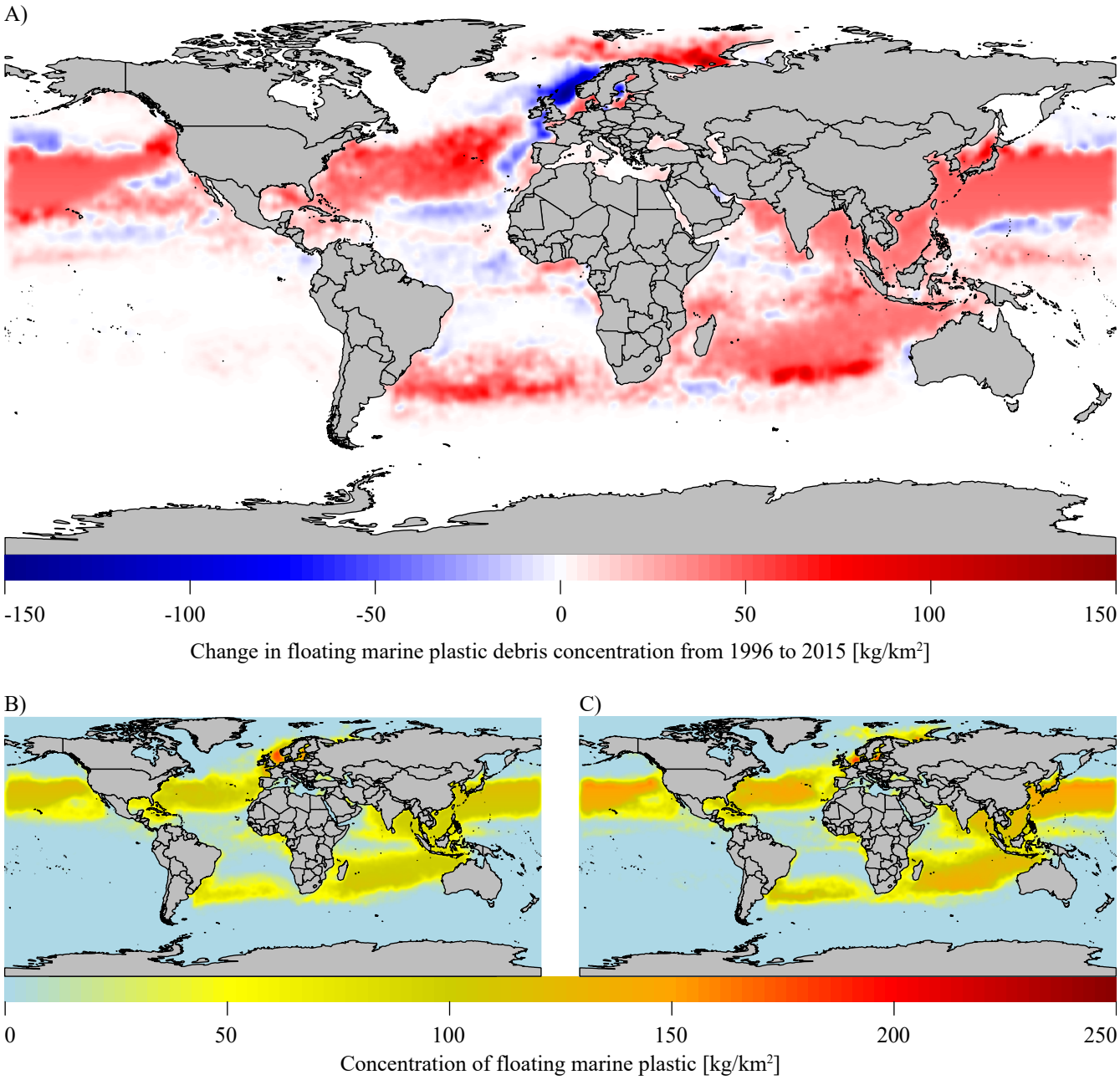

Figure 17. Results of the PARCELS simulation. A: Predicted difference in the concentrations [kg/km^2] of marine plastic debris from 1996 until 2015. B: Predicted concentrations for 1996. C: Predicted concentrations for 2015. See section 2.7 for further details.

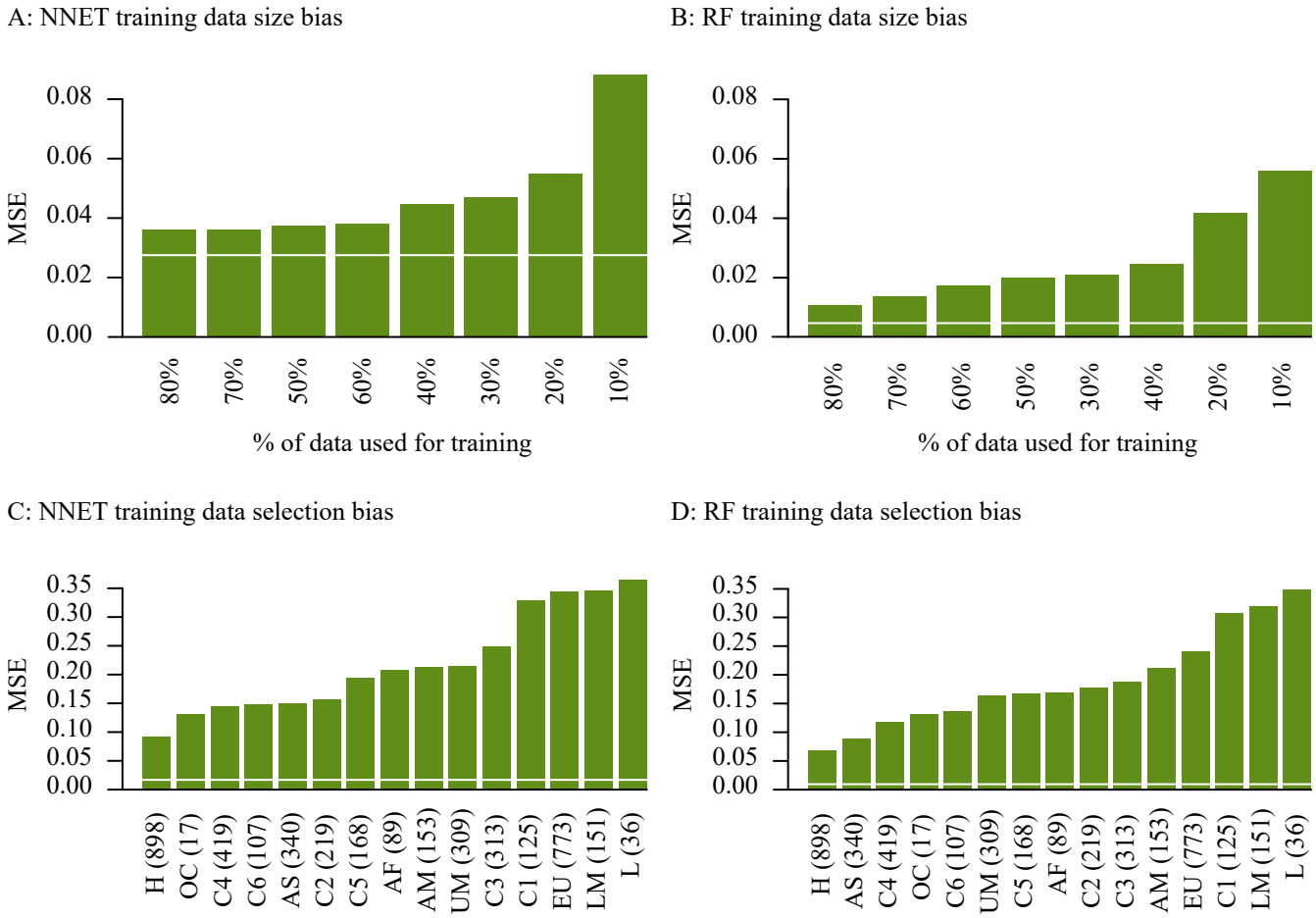

Figure 18. Results of the Model stability tests as described in section 2.4.3. A and B show the MSE for different sized training data. C and D show the MSE of the MLM trained on data selected by income, continent, or cluster. The numbers in brackets indicate how many data points were included in each training dataset (total N=1'394). The white lines indicate the MSE of the model when the training data is randomly selected and training is done on 90% of the available data. Abbreviations: H,UM,LM,L: high, upper middle, lower middle and low-income countries, OC: Oceania, AS: Asia, AF: Africa, AM:Americas, EU: Europe, C*n*: GDP-cluster *n* (compare Fig. 5).

if no changes in either policy measures or no improvements in the waste collection system occur.

4.1.3 Influence of $F_{collection}$ and the Losses from the Waste Collection System on P_O

We predict that a $F_{collection}$ of 100% would decrease the 2015 P_O by 1.7 MT, which equals 10% of the global P_O (section 3.4.2. This is mainly due to reduced plastic waste emissions in Asian and middle-income countries (1.5 Mt). This result implies that an increase of $F_{collection}$ in Asian middle-income countries can have a significant effect on the global P_O. On average, the $F_{collection}$ of all Asian countries is 65% and reaching the average level of European countries (88%) would thus result in a decrease of P_O of 0.5 Mt. For the scenario where no plastic is lost from the collection system, we predicted that the P_O could be reduced by 60% (section 3.4.2). Thereof, Asia and middle income countries contribute 90%. Thus, we predict that reducing the losses of the waste management system is the more efficient waste reduction measure compared to an increased rate of collection. As the collection of MSW is often the costliest step in the whole waste treatment process [78], the reduction of losses should be preferred over the increase of $F_{collection}$. Together, policy measures to increase waste collection rates as well as minimize losses from the waste collection system that target the Asian middle-income countries could reduce the global P_O by up to 70%. This assumes that these countries reach a 100% $F_{collection}$ and 0% losses from the waste collection system. However, already reaching levels close to European countries would reduce the global P_O significantly. Even slight improvements in these countries can therefore have a large effect on the global P_O.

4.1.4 Influence of the Transport Distances on P_O

We predict that reduced transport probabilities of plastic waste to the ocean would decrease the coastal input of P_O by 28% while the riverine input could decrease by 55% (section 3.4.2). This implies that a reduction of the riverine fraction of P_O is easier to achieve by clean-ups than a reduction of the coastal input. However, the effect of the reduced riverine load on the global P_O are marginal (3%).

4.2 Comparison of our Results with Literature Values

In this section we compare our results (as presented in the previous chapter) to estimates found in the literature. We begin with comparing our MSW_C estimates (presented in section 3.2) as predicted by the RF model. This is followed by the comparison of our P_O estimates that we presented in section 3.3.2.

4.2.1 Comparison of the MSW_C estimates

We compare or estimates of MSW_T (in t/person/year) to Jambeck et al. [4] who provide waste generation rates (in kg/person/day). Their data is based on rates of 128 countries in 2005 presented by Hoornweg&Bhada-Tata [37] and extended for a total of 192 countries. We first needed to make our 2010 estimates of MSW_T comparable to their waste generation rates. Thus, the estimates of Jambeck et al. were multiplied by 365 days/year and 1'000 kg/t to obtain t/person/year. Figure 19 shows the difference in our estimates of the waste generation rates and Jambeck et al. estimates for 58 countries where enough data was available to compare the estimates. A positive difference indicates that our prediction is higher than Jambeck et al. The mean (0.005 t/person/year) and median (-0.07 t/person/year) of the differences are close to zero. However, the estimates for some countries differ up to 2 t/person/year. We investigated the reasons for the large differences for the countries that are highlighted in Figure 19. Our RF model predicts an MSW_C that is significantly higher or lower for the following countries: Qatar (+1.0 t/person/year), Brunei Darussalam (+0.35 t/person/year), Monaco (+0.36 t/person/year), Belize (-0.87 t/person/year), and Guyana (+1.89 t/person/year). When we furthermore apply the $F_{collection}$ to estimate MSW_T, the error even increases (compare 8). While the error of the RF estimates are moderate for Qatar (6%), Brunei Darussalam (-8%) and Monaco (4%), the error of the RF estimate of Belize (-39%) can explain the difference to some extent (compare Fig. 13). No prediction error is available for Guyana. The difference of our estimates and those from Jambeck et al. for Qatar, Brunei Darussalam and Monaco can thus be explained by the difference in our available MSW_C data (obtained from [49], compare section 2.1) and the estimates of Jambeck et al. In the case of Russia and Angola, the MSW_C estimates of the RF model are close to the estimates of Jambeck et al. ($<$0.1 t/person/year difference). However, the low $F_{collection}$ of these countries (19% and 18% respectively) result in a 5-fold increased estimate of MSW_T. These $F_{collected}$ values are among the lowest of the 58 countries that we compare here. Values closer to the average $F_{collection}$ would decrease our estimated MSW_T and thus the observed difference in the estimates. Additionally, the RF estimates of Russia and Angola are also erroneous (Russia: 182%, Angola: 37%). Thus, the RF error might again explain part of the observed difference. As the error of the other countries is smaller and Jambeck et al. include estimates for countries which were not included in our model training data, this comparison gives us further confidence of the model accuracy and prediction of unseen data.

4.2.2 Comparison of the P_O estimates

Jambeck et al. [4] predict 4.8-12.7 Mt of plastic entering the ocean in 2010 while our models predicts 17.5 Mt in 2010 (table 8). However, the results of Jambeck et al. are assuming that only the population within 50 km of the coast (35.6% of the worldwide population in 2010) produce plastic waste that enters the ocean while our model calculates the waste flux to the ocean by considering the population within 100 km from the shore (44.6% in 2010). While Jambeck et al used conversion rates (15%, 25% and 40%) to estimate the fraction of this waste that enters the ocean, we predicted P_O using our WFM and distance based probabilities for P_E (section 2.5.3).

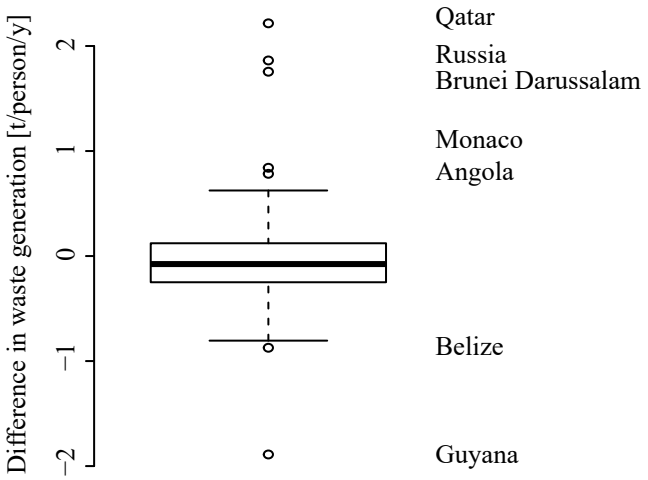

Figure 19. Difference of the MSW creation rates from Jambeck et al. [4] and our predicted MSW_C combined with averaged $MSW_{collection}$. A positive difference indicates that our prediction is higher than Jambeck et al.

A comparison is thus difficult as Jambeck et al. assume the same conversion rate for each country while our estimates include the waste management system of each country (compare Fig. 8). Additionally, we include transport by rivers (6% of P_O in 2015) which is neglected in Jambeck et al. estimates. However, down-scaling our results by the represented population (-21%, 3.5 Mt in 2010) as well as subtracting the riverine input (1.07 Mt in 2010), would change our estimate to 12.9 Mt in 2010, which accords to the highest estimate of Jambeck et al. for the same year.

Jambeck et al. furthermore provide an estimate of 31.9 Mt of mismanaged plastic waste in 2010. This represents plastic waste produced in a distance of <50 km from the shore which is not adequately disposed (defined as waste in open, uncontrolled landfills) or littered. Our estimate of 49.4 Mt of P_E in 2010 includes plastic waste that we expect to be lost from the waste collection system (27.1 Mt) or uncollected waste (22.3 Mt, compare table 7) but on a worldwide scale. It is therefore difficult to compare these estimates. Again, down-scaling our estimates by the represented population results in 39.8 Mt of inadequately disposed waste within 50 km of the shore in 2010, which is 8 Mt higher than the estimate of Jambeck et al. for the same year.

Overall, our 2010 estimates are similar to Jambeck et al. in both the quantity of plastic waste streams (Jambeck et al.: 4.8-12.7 Mt, our estimate: 17.5 Mt) as well as in the distribution among countries. Nine of our ten countries that contribute the most to P_O (table 10) ranked by their 2010 input appear in Jambeck et al. top 20 list for 2010 [4]. The only exception is Timor-Leste, for which we predict 965 kt in 2010 while it does not appear in Jambeck et al. top 20 list. While our predicted waste generation rate (0.05 t/person in 2010) is even lower than Jambeck et al. (0.29 t/person in 2010), our population is considerably higher (1.49 millions in 2010) than Jambeck et al. (0.67 millions in 2010). Additionally, Timor-Leste is among the countries with the highest fraction of MSW that is treated in dumps (49%) and only 41% of the produced MSW is collected (table A.2). Thus, our estimated impact of Timor-Leste on P_O for 2010 is higher than the 2010 estimates of Jambeck et al.

4.2.3 Comparison of Riverine Plastic Fluxes

Lebreton et al. [33] estimate that rivers annually transport 1.15 to 2.41 Mt of plastic to the oceans. They were using the MSW production estimates of Jambeck et al. for the year 2010, which we discussed in the previous section. Thus, we compare the results of Lebreton et al. to our 2010 estimates where we predict 1.07 Mt or riverine plastic input. Our estimate is thus slightly lower than the estimate of Lebreton et al. However, it has to be considered that Lebreton et al. include the size of the river catchment, drainage patterns, topography and dams, while we only consider a uniform catchment size of 20 km and a distance based probability of transport in the river (compare section 2.5.3). Thus, a comparison of the estimates is again difficult. Because Lebreton et al. used the MSW production estimates of Lebreton et al., we would expect that our estimates are larger (as our estimates of MSW production are larger than the estimates of Jambeck et al.). However, our estimates of riverine inputs are similar to Lebreton et al., which implies that our distance based probability approach results in lower estimates than the more detailed approach of Lebreton et al. Furthermore, we only include major rivers in our model (N=342) while Lebreton et al. included 40'760 watersheds worldwide. This again leads to smaller estimates of our estimates compared to Lebreton et al.

Lebreton et al. estimate that Asian rivers account for 86% of the global input while our model estimates that 69% of the global riverine input originates in Asia. However, Lebreton et al. state that Indonesian rivers account for 14.2% (200 kt) of the total global river emissions. Our model only includes four rivers in Indonesia and its fraction of the worldwide riverine input is only 0.6% (6.9 kt). This discrepancy explains the lower estimates of our model in both the total flux as well as the lower input of Asian rivers.

Schmidt et al. [34] estimate the global riverine plastic input from 0.47 Mt to 2.75 Mt. They used non linear relationships between measured plastic concentrations in rivers and the mismanaged plastic waste obtained from Jambeck et al. [4] for the year 2010. Again, our estimate for 2010 lies in the lower ranges of their prediction. Again, we explain this with the higher number of rivers used by Schmidt et al. (N=1'494 compared to our 342 rivers) as well as the lower estimates of our distance based probability approach.

For eight of our ten rivers with the highest plastic loads, our estimates are within the range of the predictions of Schmidt et al. and/or Lebreton et al. (table 12). However, we underestimate the Yangtze by at least a factor of 2 and for the Red River our estimates are three times higher. Yet, comparing the estimates of Schmidt et al. and Lebreton et al. shows that

Table 12. Comparison of our 10 rivers with the highest estimates of plastic input to the ocean in 2010 (as in table 9) with the estimates of Schmidt et al. [34] and Lebreton et al. [33].

Name	Input 2010 [t/year]		
	Our estimate	Schmidt et al. [34]	Lebreton et al. [33]
Ganges-Brahmaputra	145'691	79'075 - 12'269	172'000 - 105'000
Yellow River (Huang He)	104'282	133'810 - 19'239	
Nile	99'924	91'835 - 13'962	
Yangtze (Yangzi Jiang)	89'358	1'538'763 - 154'722	480'000 - 310'000
Indus	77'087	176'309 - 24'355	
Pearl River (Zhujiang)	62'684	57'781 - 9'400	
Red River (Yuang Jiang)	60'430	4'805 - 1'114	23'100 - 10'900
Irrawaddy	42'556	5'627 - 1'276	56'900 - 29'700
Mekong	34'997	36'761 - 6'374	37'600 - 18'800
Congo	30'202	15'052 - 2'969	

they do not agree in three out of six cases where estimates are available. This shows the high uncertainties we are still facing in riverine plastic input estimates.

4.3 Error Estimation

While we present estimates of waste production and plastic fluxes, we do not give an estimate of the error. As it has already been stated, the error of the model prediction is known from the CV but there is no estimate for the other predictions that were performed. Here we discuss possible techniques to get an error estimate.

While the prediction error of the RF and NNET models helps to see how well the model performs on the unseen test dataset, it does not provide information regarding the variance among the models. Therefore, the distribution of the country-level predictions of multiple models could be compared (e.g. one standard deviation) to get an error of the overall prediction of the NNET and RF models. This would result in an estimate of the variability of the predictions and thus the robustness of the models. Additionally to the CV error, this would also yield errors for countries where no MSW_C data is available.

A possible way to give an estimate of the errors produced in the waste flow modelling would be to create scenarios. While we do not know how accurate the data of $MSW_{collection}$ is, we could create two scenarios where the reported values are reduced or increased. This would give an idea of the influence of this parameter. The same could be done with the values of fluxes to the environment from collected and uncollected MSW. This has already be partially done with the distance where we assessed the effect of a reduced probability of transport from the shore and by rivers (compare table 11). An upper and lower estimate could be created by implementing all these scenarios together.

Even though we did not yet perform these steps to provide an overall error estimate, we will present further thoughts on how to improve the accuracy of our results in the following section.

4.4 Accuracy of the Estimates

As described above, our approach predicts MSW_C and P_O accordingly to previously published estimates. Our RF model has low prediction errors for many countries. However, for several countries, our estimated MSW_C are erroneous (compare Fig. 11 A). As our results suggest, the overall P_O estimate is largely determined by the estimate of the countries with a big influence on the global P_O. The errors of prediction of the RF model are high for several high-input countries, including China, Indonesia, Myanmar and the Philippines (compare Fig. 11 A). This produces a significant error of the global estimate. We predict that 47% of all marine plastic debris originates from China. As the RF model over-predicts China by 69%, the global plastic input in 2015 reduces by 3.6 Mt to 15.1 Mt (-19%) if the input of China is reduced by its error. Correcting the top 10 countries, who together account for 80% of the global plastic input to the oceans, results in an estimate of 14.0 Mt in 2015 (-25%). When we correct all countries by their error, the global plastic input reduces to 13.7 Mt (-27%) in 2015 and to 12.9 Mt in 2010 (-26%), which is close to Jambeck et al. [4] estimate of 4.8-12.7 Mt. This again demonstrates the high impact of the top 10 countries on the global waste flux.

Unfortunately, none of the trained RF models was able to predict these countries with an error $<$20% (compare table 6). Even tough we took the impact of each country into consideration when choosing the best model, it was thus not possible to select a model that adequately predicts the top 10 countries. This implies, that it is challenging to create a globally valid model with the available data and/or the chosen methodology. This is supported by the fact that both NNET and RF models have their lowest prediction errors in Europe while predictions for Asia are generally multiple times larger (compare Fig. 10 and 11). Also, both NNET and RF models have a lower error the higher the income of the country is. Following, we will discuss the model performance of both the RF and NNET models. We show possible sources of the errors and propose possible improvements for future projects.

4.4.1 Causes of Uncertainty

Overfitting and missing data of poor countries

There are indications that our RF model is overfitting as there are 879 branches for the 1'394 data points. The RF is thus clinging strongly to the available data and its performance on generalizing for unseen data presumably not satisfactory. Our prediction errors are low for countries where more data is available (e.g. high-income countries). As mentioned previously, out of the 1'394 data points where all PVs as well as the MSW_C were available, 898 (64%) belonged to high-income countries and 309 (22%), 151 (11%) and 36 (3%) to upper-middle, lower-middle and low-income countries respectively. However,this does not represent the global reality where 34% of all countries are rated as high-income and upper-middle, lower-middle and low-income countries account for 16%, 24% and 25% respectively. In other words, when predicting lower-middle and low-income countries, who together represent 50% of all countries, we rely on models that could only use 14% of the available data. The models could thus be biased towards high-income countries. This could help to explain why the predictions for low-income countries are generally more erroneous. We investigated the effect of training biases in he income classes (compare 18 A & B) and found that the error significantly differs among the training data subsets. Thus, an improved accuracy could be obtained by using multiple models. For both NNET and RF, one model could be trained on high-income and upper middle-income countries, where the error was similar (Fig. 10 D), and one trained on lower middle and low-income countries. For the RF, one model could be could be trained on data for each income class.

Variation of MSW_C in the PV range

Additionally to the above mentioned missing data, errors could also occur due to a limited variation of the target variable MSW_C in the range of the PVs. In Figure A.1 we show scatterplots of the PVs versus MSW_C. We marked countries with an prediction error of >30% in red. It can be seen that these countries show very little variation in MSW_C over the whole range of the PVs. The Figures A.7 and A.8 show the countries with an error in prediction of >30% (N=18) compared to the whole dataset. Often, the erroneous countries are outside the normal range where most countries accumulate. This is especially apparent in A.8 C,E&F: most of the erroneous countries have very low MSW_C values (<0.2 t/person/year) resulting in very steep slopes when compared to the PVs. The only country that does not follow this trend is Argentina. However, out of the five data points available for Argentina in the years 2008-2012, four are approximately 0.14 t/person/year while for the year 2011, 1.2 t/person/year are reported, which results in the visible spike as well as in the large error in the estimate. When neglecting this outlier, Argentina also falls in the class of erroneous countries with low reported MSW_C values of <0.2 t/person/year. The countries Angola, Belize and Russia each have 2 data points and each country has one reported MSW_C value of >0.2 t/person/year.

Besides having low reported MSW_C values, the erroneous countries are also accentuated in their PV range (Fig. A.7 and A.8 C,E&F). In the three-dimensional space, many of the erroneous countries have a high fraction of young population coupled with a low fraction of urban population and sometimes strong greenhouse gas production and are thus outside the normal range of most countries. When comparing the slopes in Figure A.8 A, several of the erroneous countries show a slope that differs compared to the other countries. In these countries, a fast rate of urbanization meets either a decrease in birth rates or a prolonged life expectancy. These fundamental changes in the socio-economic situation of the countries might be a reason why a prediction is difficult. Especially China shows a rapid urbanization rate whilst its PV range is often less extreme than for the other erroneous countries. Furthermore, several countries show strong fluctuations in the PV *greenhouse gas* which seems to have less effect on the prediction performance. Thus, many of the erroneous countries have values in the PVs that are outside the range of most other countries and all of them have low reported MSW_C values. However, the countries Argentina, Russia and Ukraine appear in the normal range in plot A.7. While Argentina, as already stated, has an outlier in its data, the Ukraine has an unusual combination of a low fraction of young population and low MSW_C (compare Fig. A.8 C) which might explain the error in prediction. However, the error of Russia cannot be explained with unusual values in its PVs. Here, a possible explanation lies in the data availability. As Russia has only two reported MSW_C values, one for 2000 (0.30 t/person/year) and one for 2012 (0.07 t/person/year) it is possible that the data was collected using different methodologies.

Data quality

As mentioned above, the initial data includes several countries with spikes or "jumps" (e.g. Argentina or Russia). Hoornweg&Bhada-Tata [37] note, that the MSW data should be considered with caution as data collection methodologies and definitions contain inconsistencies. Thus, the MSW_C data we used to train might have been collected slightly differently among the countries. Additionally, several countries do not provide continuous data over several years and the information could thus be collected at a non-representative moment [37]. The same issue has to be considered for each of our variables used in the MSW production modelling or the waste flow modelling. Generally, the data completeness is highest for high-income countries and the consistency is thus higher as outliers can be recognized. Additionally, the data for many low and middle-income countries is further compromised due to limited technological possibilities to accurately measure MSW_C [37] or lacking funds to accurately characterize the waste composition and characteristics. Thus, these countries often provide estimates rather than actual measurements [37]. Furthermore, a changing economic situation in low and middle-income countries often leads to large variations in waste quantities which not necessarily represent

increased waste production but a changed collection rate due to increased values of recoverable secondary materials in the waste (e.g. metals or high-value plastics) [37].

Small scale effects
Besides the data quality, effects on a local scale could also effect the model prediction performance. Possible income gaps between rural and urban population or between regions within a country are not represented in our models. Neither do we include topography, weather and climate, locations of known large deposition sites nor data on river run-off or dams, which could possibly have severe effects on the local conditions and thus the release of plastic waste to the environment. Thus, using country-averaged data might lead to a discrepancy in MSW_C within countries. This effect possibly affects large or highly populated countries more than small and sparsely populated countries. However, as data availability is already an issue on country level, including information on sub-national regions is especially difficult for socio-economic factors that cannot be measured using remote sensing. However, geographical and climatological data is available and could be used to get more detailed waste flows. This has already been done by Lebreton et al. [33] who used hydrological data coupled with the location of dams to estimate riverine inputs.

Missing information of the waste flow
As stated in section 2.5.2 and figure 9, the losses from the collected and uncollected waste to the natural environment are unknown for most countries. For the middle-income countries we used the known fluxes from China and the Philippines and scaled each country with its waste management system compared to the reference countries. However, for high and low-income countries, no data was available and we estimated these fluxes assuming a linear trend of losses and income. This could possibly result in erroneous estimates for high and low-income countries. However, our models estimate that only 10% of the global input to the ocean originates from these countries (compare table A.4). Thus, the uncertainty in the losses does only marginally affect the global estimates and the large error in predicting the MSW_C in China has a much larger effect. Nonetheless would literature based loss estimates further increase the confidence in our estimates.

Predictor variable selection
The three variables selected to predict the MSW (*urban population*, *greenhouse gas* and *young population*) were important in all analyses (compare table 5). When comparing the correlations between these variables (Fig. 4), it can be seen that they are not strongly correlated using either the Pearson or Spearman correlations. Additionally, the correlations of each selected PV to the non-selected PVs is often stronger. Furthermore, the selected variables are often almost complete and almost no extrapolation is needed (compare table A.1). As the the different selection processes agreed on the same three PVs and their data is often available without interpolation, we expect only marginal uncertainty arising in the models due to the selected models. However, as the scatterplots of MSW_C versus the PVs (Fig. A.1 shows, the explanatory power for several countries is low. Furthermore, it might be that the $MSW_{collection}$ in low-income countries is governed by other factors than our seven PVs. As it has already been noted in section 2.4.2, parameters such as policy factors or public attitudes that are known to affect MSW_C are not included in our PV set. Thus, the model performance in low-income countries could maybe be increased by adding such parameters.

Bias in the training data
As Figure 18 shows, training the models on a selected subset can have a huge effect. Surprisingly, training solely on data from European countries triples the error compared to training on high-income countries (Fig. 18 A&B). As 34 of the total 45 European countries are classified high-income, we expected similar error for Europe and high-income countries. This might be caused by autocorrelation as the time series data of European countries was mostly complete (compare tableA.2. Additionally to the 34 European countries, the 59 countries classified as high-income include 14 Asian, three Oceanian, one African, four Northern American, and two South American countries and the Caribbean islands. Thus, as this training dataset includes all continents and 5 out of 7 GDP clusters (compare Fig. 5), the diversity might explain why training on high-income countries yields the best results. On the other hand, the 28 countries classified as low-income only include 26 African and two Asian countries, that are all in the cluster 1 and the dataset is thus more uniform, which might explain the tripled error compared to high-income country trained models. However, training on Oceanian countries results also in relatively small errors despite the small dataset (17 points) and low diversity (only three countries) resulting in an uniform training dataset.
To minimize the effect of biases in the training data, we trained 10'000 models on randomly selected training and test data and performed a 10-fold CV. A stricter selection in the CV might prefer models with a higher prediction power for unseen data. However, if this significantly lowers the prediction error needs to be tested.

4.5 Limitations of the Model

While the above mentioned measures could help to improve our estimates, we now want to highlight the main limitations of our model.
Our approach predicts the waste production on a 0.1x0.1° global raster. We include several assumptions that do not fully represent the reality. First of all we assume that the waste is produced where people live. This completely ignores the effect of tourism, where waste is produced away from home. Thus, the waste production in popular tourist destinations might be underestimated by our approach. Tourism has been found to increase the marine litter in the Mediterranean Sea by +40% in the summer caused by >200 million tourists visiting

the areas each year [7]. Similar effects might be observed in other popular tourist destinations. Including tourism as a variable in the model might help to account for this fact.
Furthermore, we assume that the waste produced in each cell is also treated in the same location and we do not account for the trade of waste and recycling material. This results in two limitations of our model. First, this might result in erroneous predictions of fluxes to the oceans. If waste produced far from the shore is transported to a dump site closer to the sea, its chances of reaching the ocean rise. Similarly, waste produced in large cities at the shore of rivers might be discarded outside the city's boundaries further from the river and with a lower chance of transport. Such effects are not included in our model. Secondly, if waste is exported to another country, the treatment will occur in a system that is different than in the original country. The European Union accounted for 34% of the global export of non-hazardous waste in 2014, of which 60% are transported to non-OECD countries of which 31% went to China [79]. On the other hand, the European Union also accounts for 20% of the global imports of non-hazardous waste. If the trade of waste has an influence on the losses to the environment and thus influxes to the ocean, it is not represented in our models.

4.6 Evaluations of the Goals of this Thesis

Here we reflect if the goals presented in the introduction have been reached. Our first hypothesis is that *modelling MSW_C is possible using the PVs GDP, share of urban population, fraction of population aged 0-14 years, fraction of households with access to electricity, energy consumption and greenhouse gas emissions.*
Using the PVs *urban population*, *energy consumption*, and *young population* we were able to predict MSW_C for 173 countries. We compared our results to the estimates of Jambeck et al. [4]) and found similar results. However, the model performance is significantly better the higher the country's income is. Especially countries classified as low-income were erroneous. This can to some degree be explained by the low variance of the PVs for these countries. However, the NNET and RF model might not be ideal with the data we used. Simpler models such as multiple linear regression could result in less error-prone estimates. Creating multiple models and focusing on countries with a high impact on P_O might also help to improve the overall error of the predictions.
Our second hypothesis is that *trends in the global production and disposal of MSW are expected to affect the plastic waste that enters the environment*. The increase in MSW production can be seen in our prediction of P_E and P_O. We predict an increase of P_E from 31.8 Mt in 1990 to 51.0 Mt in 2015 (+60%) and a growth of P_O from 13.4 Mt in 1990 to 18.7 Mt in 2015 (+40%). The faster growth of P_E compared to P_O can be explained by an increase of plastic waste production in regions without direct access to the sea. P_E in these regions can only reach the ocean by riverine transport.
Our third hypothesis is that *the income class of each country is expected to influence the amount of plastic waste due to decreased effectiveness of the waste collection system with decreasing income*. Our estimates of P_E and P_O are strongly influenced by the income. We predict that middle-income countries account for 90% of P_O while representing only 58% of the world's population. Using a scenario where all waste is collected we could show that P_O would sink by 10%. Again, the effect of middle-income countries accounts for 90% of this reduction. However, a scenario where no waste is lost from the waste collection system predicted even higher decreases in P_O (59%). Thus, the effect of the income class on P_O is manly explained due to losses from the waste treatment system and not due to the fraction of waste collection.
Our fourth hypothesis is that *advecting our predicted plastic fluxes to the ocean in an Lagrangian particle advection model will result in similar plastic concentration patterns as in the observed aggregation zones*. We could show that the PARCELS model using the Eulerian total current and our predicted plastic inputs to the ocean is able to recreate the known aggregation zones. However, comparing these results with measurements is difficult because we did not include any degradation processes on the plastic particles. These can strongly affect the behaviour of floating plastic debris. Thus, our simulation could be used as benchmark to compare the relative effects of changing plastic inputs on the spatio-temporal level.

In the following chapter we will provide a short conclusion of the main findings of this thesis and provide further research questions.

5. Conclusions

With an increasing production of plastic goods, the environmental concerns are rapidly growing. Plastic debris in the ocean is a known threat for marine life and causes economic loss. We predict an increase in total plastic debris entering the ocean (P_O) from 13.4 Mt in 1990 to 18.7 Mt in 2015 (+40%). The ten countries with the highest impact on P_O account for 15.1 Mt in 2015 which represents 81% of the global P_O. We differ between inputs from riverine and coastal sources that account for 6% and 94% of P_O respectively. In the time period 1990-2015, the riverine input has more than doubled (+134%). Geographically, the strongest increase in P_O can be seen in Asia (+4.3 Mt) and economically, middle-income countries (+4.8 Mt) show the strongest increase. In 2015, 84% of P_O (15.7 Mt) originated from Asia and 90% of P_O (16.8 Mt) originated from middle-income countries (representing 58% of the world's population). We predict that from the 18.7 Mt of P_O in 2015, 60% (11.2 Mt) originate from losses in the waste collection system and 40% (7.5 Mt) originate from waste that was not collected. Thus, measures which target an improved waste collection and treatment system in Asia have the potential to reduce the global P_O by up to 8.4 Mt (45% of P_O) in an ideal scenario. Already minor improvements could have a significant impact on the global P_O. These results imply that a reduction of plastic waste is a more effective measure than ocean clean-ups to reduce P_O. Besides the period 1990-2015, we also created future estimates for 2030 and 2050. We predict that the waste production of African and Asian countries will increase, which will also further increase P_O to 19.7 Mt in 2030 and 21.6 Mt in 2050. Besides P_O we also estimated the plastic waste entering the environment (P_E). We predict an increase from 31.8 Mt in 1990 to 51.0 Mt in 2015 (+60%). The faster growth of P_E compared to P_O can be explained by an increase of plastic waste production in regions without direct access to the sea. P_E in these regions can only reach the ocean by riverine transport.

We used a NNET and RF model to predict the amounts of produced municipal solid waste using the predictor variables *urban population*, *greenhouse gas* and *young population*. The RF model has an error of prediction of 2% for high-income countries. This error increases for middle (12% for upper and 16% for lower) and low-income countries (86%). The NNET model is more erroneous in 103 out of 106 countries and shows the same trend regarding prediction error and the country's income as the RF model.

High errors might arise due to the quality of the predictor variables, missing data, biases in the training data or overfitting of the model. To increase the overall prediction accuracy, we propose to create multiple models for the income classes and to account for autocorrelation in the time-series data. Furthermore, a choice of less complex models (e.g. multiple linear regression) could help to increase the prediction performance. A focus to reduce the prediction error should be laid on the countries that contribute most to P_O. Additionally, we want to highlight the need for more reliable data on the quantity and composition of each country's waste, as well as of the composition of the waste management system, the estimated waste collection rates and losses from the collection system. The recently published report *"What a waste 2.0: A global Snapshot of Solid Waste Management to 2050"* [78] (published on the 20th September 2018 and thus not included here) provides an updated and extended version of our data and might be useful to obtain such data.

We were able to show that modelling the waste production of each country in the time period 1990-2015 is possible using socio-economic predictors. While the errors of the estimates could be reduced by increasing the number of models (e.g. for each income class) and modelling techniques (e.g. multiple linear regression), our results are still comparable to similar studies. We show that reducing losses from the waste treatment system and increasing the fraction of waste collection in Asian countries could significantly reduce the plastic entering the ocean. Our results help to gain further knowledge on the spatio-temporal distribution of P_O. This helps to better understand the *missing plastic problem* as well as to better predict and quantify impacts of plastic waste on both the marine environment and the economy (e.g. tourism or the fishing industry). Furthermore, the results improve our understanding of waste flows from the municipal level via the waste collection and treatment system to the environment and, finally, to the oceans. Understanding the waste flows is a crucial step to determine efficient actions of P_E and P_O prevention. This can be help to reduce P_O by target-oriented measures or sanctions on specific parts of the waste flow (e.g. collection rates or losses). Additionally, our work will help to create improved models to predict the waste production or P_E and P_O. However, there are still various open questions. A better knowledge on the drivers of municipal solid waste generation could help to better predict the production of waste. Furthermore, more detailed insights on the waste collection and treatment on a national and regional level cold help to better understand the pathways of (plastic) waste and pinpoint losses from the system. Both could help to reduce the plastic waste that enters the environment and the oceans in the future. Additionally, a better understanding of the degradation processes of plastic in the marine environment can help to understand the fate and sinks of the marine plastic debris that already entered our oceans. Both prevention and mitigation of marine plastic debris are essential to protect the marine environment from the deleterious effects of plastic waste.

6. Acknowledgements

I would like to thank Dr. Meike Vogt and Dr. Charlotte Laufkötter for their supervision of this thesis. Their expert advice, profound criticism, support and encouragement throughout the process was extremely helpful. Furthermore I wish to thank Dr. Fabio Benedetti for his help on both the data analysis and model selection process as well as for the fruitful discussions on the modelling approaches.

References

[1] Roland Geyer, Jenna R Jambeck, and Kara Lavender Law. "Production, use, and fate of all plastics ever made". In: *Science Advances* 3.7 (2017), pp. 1–5. ISSN: 23752548. DOI: 10.1126/sciadv.1700782.

[2] Anthony L. Andrady. "Microplastics in the marine environment". In: *Marine Pollution Bulletin* 62.8 (2011), pp. 1596–1605. ISSN: 0025326X. DOI: 10.1016/j.marpolbul.2011.05.030.

[3] Richard C. Thompson et al. "Lost at Sea: Where Is All the Plastic?" In: *Science* 304.5672 (2004), p. 838. ISSN: 00368075. DOI: 10.1126/science.1094559.

[4] Jenna R. Jambeck et al. "Plastic waste inputs from land into the ocean". In: *Science* 347.6223 (2015), pp. 768–771. ISSN: 10959203. DOI: 10.1126/science.1260352.

[5] Edward J Carpenter and K L Smith. "Plastics on the Sargasso Sea Surface". In: *Science* 175.4027 (1972), p. 12401241. DOI: 10.1126/science.175.4027.1240.

[6] Merel Kooi et al. "Modeling the fate and transport of plastic debris in freshwaters: Review and Guidance". In: *Freshwater Microplastics*. Ed. by M Wagner and S Lambert. Vol. 58. Cham, Switzerland: Springer, 2018, pp. 125–152. DOI: 10.1007/978-3-319-61615-5_7.

[7] Eva Alessi and Guiseppe Di Carlo. *Out of the Plastic Trap Saving the mediterranean from plastic pollution*. Tech. rep. Rome, Italy: World Wide Found For Nature (WWF), 2018, pp. 1–28.

[8] Susanne Kühn, Elisa L Bravo Rebolledo, and J. A. Van Franeker. "Deleterious Effects of Litter on Marine Life". In: *Marine Anthropogenic Litter*. Ed. by M Bergmann, L Gutow, and M Klages. 1st ed. Cham, Switzerland: Springer, 2015. Chap. 4, pp. 75–116. ISBN: 978-3-319-16510-3.

[9] HJ Auman et al. "Plastic ingestion by Laysan albatross chicks on Sand Island, Midway Atoll, in 1994 and 1995". In: *Albatross biology and Conservation* 1 (1997), pp. 239–244.

[10] Matthew Cole et al. "The impact of polystyrene microplastics on feeding, function and fecundity in the marine copepod Calanus helgolandicus". In: *Environmental Science and Technology* 49.2 (2015), pp. 1130–1137. ISSN: 15205851. DOI: 10.1021/es504525u.

[11] Patrick M. Canniff and Tham C. Hoang. "Microplastic ingestion by Daphnia magna and its enhancement on algal growth". In: *Science of the Total Environment* 633 (2018), pp. 500–507. ISSN: 18791026. DOI: 10.1016/j.scitotenv.2018.03.176.

[12] Ellen Besseling et al. "Nanoplastic affects growth of S. obliquus and reproduction of D. magna". In: *Environmental Science and Technology* 48.20 (2014), pp. 12336–12343. ISSN: 15205851. DOI: 10.1021/es503001d.

[13] Dennis Brennecke et al. "Ingested microplastics are translocated to organs of the tropical fiddler crab Uca rapax". In: *Marine Pollution Bulletin* 96.1-2 (2015), pp. 491–495. ISSN: 18793363. DOI: 10.1016/j.marpolbul.2015.05.001.

[14] Concepción Martínez-Gómez et al. "The adverse effects of virgin microplastics on the fertilization and larval development of sea urchins". In: *Marine Environmental Research* 130 (2017), pp. 69–76. ISSN: 18790291. DOI: 10.1016/j.marenvres.2017.06.016.

[15] Stephanie L. Wright, Richard C. Thompson, and Tamara S. Galloway. "The physical impacts of microplastics on marine organisms: A review". In: *Environmental Pollution* 178 (2013), pp. 483–492. ISSN: 02697491. DOI: 10.1016/j.envpol.2013.02.031.

[16] Jessica Reichert et al. "Responses of reef building corals to microplastic exposure". In: *Environmental Pollution* 237 (2017), pp. 955–960. ISSN: 18736424. DOI: 10.1016/j.envpol.2017.11.006.

[17] Chelsea M. Rochman. "The Complex Mixture, Fate and Toxicity of Chemicals Associated with Plastic Debris in the Marine Environment". In: *Marine Anthropogenic Litter*. Ed. by M Bergmann, L Gutow, and M Klages. 1st ed. Cham, Switzerland: Springer, 2015. Chap. 5, pp. 117–140. ISBN: 978-3-319-16510-3.

[18] Chelsea M. Rochman et al. "Classify plastic waste as hazardous". In: *Nature* 494 (2013), pp. 169–171. DOI: 10.1038/494169a.

[19] Joana C. Prata et al. "Influence of microplastics on the toxicity of the pharmaceuticals procainamide and doxycycline on the marine microalgae Tetraselmis chuii". In: *Aquatic Toxicology* 197.February (2018), pp. 143–152. ISSN: 18791514. DOI: 10.1016/j.aquatox.2018.02.015.

[20] Qiqing Chen et al. "Pollutants in Plastics within the North Pacific Subtropical Gyre". In: *Environmental Science and Technology* 52.2 (2018), pp. 446–456. ISSN: 15205851. DOI: 10.1021/acs.est.7b04682.

[21] Frederic Gallo et al. "Marine litter plastics and microplastics and their toxic chemicals components: the need for urgent preventive measures". In: *Environmental Sciences Europe* 30.13 (2018). ISSN: 21904715. DOI: 10.1186/s12302-018-0139-z. URL: https://doi.org/10.1186/s12302-018-0139-z.

[22] Albert A. Koelmans et al. "Microplastic as a Vector for Chemicals in the Aquatic Environment: Critical Review and Model-Supported Reinterpretation of Empirical Studies". In: *Environmental Science and Technology* 50.7 (2016), pp. 3315–3326. ISSN: 15205851. DOI: 10.1021/acs.est.5b06069.

[23] European Comission. *Single-use plastics : New EU rules to reduce marine litter*. Brussels, Belgium, 2018.

[24] Judi W. Wakhungu. *The Environmental Management and Co-ordination Act Cap. 387*. Nairobi, Kenya, 2017.

[25] A. Cozar et al. "Plastic debris in the open ocean". In: *PNAS* 111.28 (2014), pp. 10239–10244. ISSN: 0027-8424. DOI: 10.1073/pnas.1314705111.

[26] Tim Kiessling, Lars Gutow, and Martin Thiel. "Marine Litter as Habitat and Dispersal Vector". In: *Marine Anthropogenic Litter*. Ed. by Melanie Bergmann, Lars Gutow, and Michael Klages. 1st ed. Cham, Switzerland: Springer, 2015. Chap. 6, pp. 141–181. ISBN: 978-3-319-16510-3.

[27] Richard C Thompson. "Microplastics in the Marine Environment: Sources, Consequences and Solutions". In: *Marine Anthropogenic Litter*. Ed. by M Bergmann, L Gutow, and M Klages. 1st ed. Cham, Switzerland: Springer, 2015. Chap. 7, pp. 185–200. ISBN: 978-3-319-16510-3.

[28] Marcus Eriksen et al. "Plastic Pollution in the World's Oceans: More than 5 Trillion Plastic Pieces Weighing over 250,000 Tons Afloat at Sea". In: *PLoS ONE* 9.12 (2014), pp. 1–15. ISSN: 19326203. DOI: 10.1371/journal.pone.0111913.

[29] Erik Van Sebille et al. "A global inventory of small fl oating plastic debris". In: *Environmental Research Letters* 10.12 (2015), pp. 1–11. ISSN: 1748-9326. DOI: 10.1088/1748-9326/10/12/124006.

[30] Atsuhiko Isobe et al. "Microplastics in the Southern Ocean". In: *Marine Pollution Bulletin* 114.1 (2017), pp. 623–626. ISSN: 18793363. DOI: 10.1016/j.marpolbul.2016.09.037.

[31] C. J. Moore et al. "A comparison of plastic and plankton in the North Pacific Central Gyre". In: *Marine Pollution Bulletin* 42.12 (2001), pp. 1297–1300. ISSN: 0025326X. DOI: 10.1016/S0025-326X(01)00114-X.

[32] L. Lebreton et al. "Evidence that the Great Pacific Garbage Patch is rapidly accumulating plastic". In: *Scientific Reports* 8.4666 (2018), pp. 1–15. ISSN: 2045-2322. DOI: 10.1038/s41598-018-22939-w.

[33] Laurent C.M. Lebreton et al. "River plastic emissions to the world's oceans". In: *Nature Communications* 8 (2017), pp. 1–10. ISSN: 20411723. DOI: 10.1038/ncomms15611.

[34] Christian Schmidt, Tobias Krauth, and Stephan Wagner. "Export of Plastic Debris by Rivers into the Sea". In: *Environmental Science & Technology* 51 (2018), pp. 12246–12253. DOI: 10.1021/acs.est.7b02368.

[35] Daniel Cressey. "The Plastic Ocean". In: *Nature* 536 (2016), pp. 263–265. DOI: 10.1038/536263a.

[36] Jan A. Van Franeker and Kara Lavender Law. "Seabirds, gyres and global trends in plastic pollution". In: *Environmental Pollution* 203 (2015), pp. 89–96. ISSN: 18736424. DOI: 10.1016/j.envpol.2015.02.034.

[37] Daniel Hoornweg and Perinaz Bhada-Tata. *What a Waste: A Global Review of Solid Waste Management*. Tech. rep. Urban Development Series No. 15. Washington DC, USA: The World Bank, 2012, pp. 1–101.

[38] F. Galgani, G. Hanke, and T. Maes. "Global distribution, composition and abundance of marine litter". In: *Marine Anthropogenic Litter*. Ed. by M Bergmann, L Gutow, and M Klages. 1st ed. Cham, Switzerland: Springer, 2015. Chap. 2, pp. 29–56. ISBN: 978-3-319-16510-3.

[39] National Oceanic and Atmospheric Administration, Woods Hole Sea Grant, and Oliver Lüde. *How long until it's gone? Estimated decomposition rates of common marine debris items*. 2014.

[40] Amanda L. Dawson et al. "Turning microplastics into nanoplastics through digestive fragmentation by Antarctic krill". In: *Nature Communications* 9.1 (2018), pp. 1–8. ISSN: 20411723. DOI: 10.1038/s41467-018-03465-9.

[41] Chris Sherringtn. *Plastics in the Marine Environment - Eunomia*. 2016. DOI: 10.1146/annurev-marine-010816-060409. URL: http://www.annualreviews.org.

[42] Albert A. Koelmans et al. "All is not lost: Deriving a top-down mass budget of plastic at sea". In: *Environmental Research Letters* 12.11 (2017), pp. 1–9. ISSN: 17489326. DOI: 10.1088/1748-9326/aa9500.

[43] K. Brunner et al. "Passive buoyant tracers in the ocean surface boundary layer: 2. Observations and simulations of microplastic marine debris". In: *Journal of Geophysical Research: Oceans* 120 (2015), pp. 7559–7573. ISSN: 21699275. DOI: 10.1002/2015JC010840.

[44] Katia Karousakis. "The Economics and Policy of Municipal Soled Waste Management". Doctoral Thesis. Ann Arbor, USA: University College London, 2006, pp. 1–238. URL: http://discovery.ucl.ac.uk/id/eprint/1444764.

[45] Katia Karousakis. "The drivers of MSW generation, disposal and recycling: Examining OECD inter-country differences". In: *Waste and Environmental Policy*. Ed. by Massimiliano Mazzanti and Anna Montini. 1st ed. Abingdon, UK: Routledge, 2009. Chap. 4, pp. 91–104. DOI: 10.4324/9780203881378.

[46] IPCC. *Climate Change 2014: Synthesis Report*. Geneva, Switzerland: IPCC, 2014, pp. 1–151. DOI: 10.1017/CBO9781107415324.

[47] Daniel Hoornweg, Perinaz Bhada-Tata, and Christopher Kennedy. "Peak Waste: When Is It Likely to Occur?" In: *Journal of Industrial Ecology* 19.1 (2015), pp. 117–128. ISSN: 15309290. DOI: 10.1111/jiec.12165.

[48] United Nations Environment Programme. *Global Waste Management Outlook*. Tech. rep. Nairobi, Kenya: United Nations Environment Programme, 2015, pp. 1–334.

[49] United Nations Department of Economic and Social Affairs: Statistics Division. *Total amount of municipal waste collected*. 2018. URL: https://unstats.un.org/unsd/envstats/qindicators.cshtml (visited on 04/06/2018).

[50] United Nations Department of Economic and Social Affairs: Population Division. *World Urbanization Prospects: The 2014 Revision*. 2014. URL: https://esa.un.org/unpd/wpp/ (visited on 04/06/2018).

[51] United Nations Department of Economic and Social Affairs: Population Division. *World Urbanization Prospects: The 2014 Revision*. 2014. URL: https://esa.un.org/unpd/wup/ (visited on 04/06/2018).

[52] The World Bank. *Population ages 0-14, total*. 2014. URL: https://data.worldbank.org/indicator/SP.POP.0014.TO (visited on 04/06/2018).

[53] The World Bank. *Access to electricity (% of population)*. 2018. URL: https://data.worldbank.org/indicator/EG.ELC.ACCS.ZS (visited on 04/06/2018).

[54] The World Bank. *GDP (current US$)*. 2018. URL: https://data.worldbank.org/indicator/NY.GDP.MKTP.CD (visited on 04/06/2018).

[55] The World Bank. *Energy use (kg of oil equivalent per capita)*. 2018. URL: https://data.worldbank.org/indicator/EG.USE.PCAP.KG.OE (visited on 04/06/2018).

[56] The World Bank. *Total greenhouse gas emissions (kt of CO2 equivalent)*. 2018. URL: https://data.worldbank.org/indicator/EN.ATM.GHGT.KT.CE (visited on 04/06/2018).

[57] United Nations Department of Economic and Social Affairs: Statistics Division. *Municipal Waste Treatment*. 2016. URL: https://unstats.un.org/unsd/envstats/qindicators.cshtml (visited on 04/06/2018).

[58] The World Bank. *GNI (current US$)*. 2018. URL: https://data.worldbank.org/indicator/NY.GNP.MKTP.CD (visited on 04/06/2018).

[59] United Nations Department of Economic and Social Affairs: Statistics Division. *Total population served by municipal waste collection*. 2018. URL: https://unstats.un.org/unsd/envstats/qindicators.cshtml (visited on 04/06/2018).

[60] Roger Bivand et al. *Bindings for the 'Geospatial' Data Abstraction Library*. 2018. URL: https://cran.r-project.org/web/packages/rgdal/index.html.

[61] Robert J. Hijmans et al. *Geographic Data Analysis and Modeling*. 2017. URL: https://cran.r-project.org/web/packages/raster/index.html.

[62] GlobCurrent Consortium. *GlobCurrent Product Data Handbook: The Combined Geostrophy + Ekman Currents*. 2016. URL: http://www.globcurrent.org/.

[63] OceanDataLab. *GlobCurrent portal*. 2018. URL: https://globcurrent.oceandatalab.com (visited on 07/06/2018).

[64] Socioeconomic Data and Applications Center. *Gridded Population of the World V3*. 2005. URL: http://sedac.ciesin.columbia.edu/data/set/gpw-v3-population-density (visited on 04/24/2018).

[65] Socioeconomic Data and Applications Center. *Gridded Population of the World V4*. 2017. URL: http://sedac.ciesin.columbia.edu/data/set/gpw-v4-population-density-rev10 (visited on 04/24/2018).

[66] Natural Earth. *Rivers + lake centerlines V4.0.0*. URL: http://www.naturalearthdata.com/downloads/10m-physical-vectors/10m-rivers-lake-centerlines/ (visited on 04/24/2018).

[67] GlobCurrent Consortium. *Total HS currents*. 2017. URL: ftp://eftp.ifremer.fr/data/globcurrent/v3.0/global%7B%5C_%7D025%7B%5C_%7Ddeg/total%7B%5C_%7Dhs/.

[68] Christophe Genolini and Elie Guichard. *K-Means for Longitudinal Data using Shape-Respecting Distance*. 2016. URL: https://cran.r-project.org/web/packages/kmlShape/index.html.

[69] Stacy L. Özesmi, Can O. Tan, and Uygar Özesmi. "Methodological issues in building, training, and testing artificial neural networks in ecological applications". In: *Ecological Modelling* 195.1-2 (2006), pp. 83–93. ISSN: 03043800. DOI: 10.1016/j.ecolmodel.2005.11.012.

[70] Stefan Fritsch et al. *Training of neural networks*. 2016. URL: https://cran.r-project.org/web/packages/neuralnet/index.html.

[71] Leo Breiman et al. *Breiman and Cutler's Random Forests for Classification and Regression*. 2018. URL: https://cran.r-project.org/web/packages/randomForest/index.html.

[72] Vladimir Svetnik et al. "Random Forest: A Classification and Regression Tool for Compound Classification and QSAR Modeling". In: *Journal of Chemical Information and Computer Sciences* 43.6 (2003), pp. 1947–1958. ISSN: 00952338. DOI: 10.1021/ci034160g.

[73] McKinsey Center for Business and Environment and Ocean Conservancy. *Stemming the Tide. Land-based strategies for a plastic - free ocean*. Tech. rep. McKinsey Center for Business and Environment, 2015, pp. 1–48. URL: http://www.oceanconservancy.org/our-work/marine-debris/stop-plastic-trash-2015.html.

[74] Skye Morét-Ferguson et al. "The size, mass, and composition of plastic debris in the western North Atlantic Ocean". In: *Marine Pollution Bulletin* 60.10 (2010), pp. 1873–1878. ISSN: 0025326X. DOI: 10.1016/j.marpolbul.2010.07.020.

[75] Rachel Hurley, Jamie Woodward, and James J. Rothwell. "Microplastic contamination of river beds significantly reduced by catchment-wide flooding". In: *Nature Geoscience* 11 (2018), pp. 251–257. ISSN: 17520908. DOI: 10.1038/s41561-018-0080-1.

[76] Michael Lange and Erik Van Sebille. "Parcels v0.9: Prototyping a Lagrangian ocean analysis framework for the petascale age". In: *Geoscientific Model Development* 10.11 (2017), pp. 4175–4186. ISSN: 19919603. DOI: 10.5194/gmd-10-4175-2017.

[77] Brian Ripley et al. *Support Functions and Datasets for Venables and Pipley's MASS*. 2018. URL: https://cran.r-project.org/web/packages/MASS/index.html.

[78] Silpa Kaza et al. *What A Waste 2.0: A Global Snapshot on Solid Waste Management to 2050.* Tech. rep. Urban Development Series. Washington DC: The World Bank, 2018, pp. 1–296. DOI: 10.1596/978-1-4648-1329-0.

[79] European Commission. *Waste Shipment*. 2015. URL: http://ec.europa.eu/trade/import-and-export-rules/export-from-eu/waste-shipment/ (visited on 09/12/2018).

[80] Ellen MacArthur Foundation, World Economic Forum, and McKinsey & Company. *The New Plastics Economy: Rethinking the future of plastics*. Tech. rep. Davos, Switzerland: Ellen MacArthur Foundation, 2016, pp. 1–36. DOI: 10.1103/Physrevb.74.035409.

[81] PlasticsEurope. *Plastics – the Facts 2016*. Tech. rep. Brussels, Belgium: PlasticsEurope, 2016, pp. 1–38.

[82] PlasticsEurope. *Plastics – the Facts 2017*. Tech. rep. Brussels, Belgium: PlasticsEurope, 2017, pp. 1–44.

7. Appendix

7.1 Abbreviations

$F_{collection}$	The fraction of MSW_T that is collected by a waste collection system.
F_{dumps}	Fraction of MSW_C that is disposed in dumps.
$F_{landfills}$	Fraction of MSW_C that is disposed in landfills.
$F_{plastic}$	Fraction of plastic in MSW_C.
$F_{recycled}$	Fraction of MSW_C that is recycled.
MLM	Machine learning models used to estimate MSW_C.
MSW	Municipal solid waste.
MSW_C	Amount of MSW that is collected by a waste collection system.
MSW_T	Total amount of MSW that is produced.
NNET	Neural network model used to predict MSW_C.
P_E	Amount of plastic that leaked into the environment.
P_O	Amount of plastic that leaks into the oceans.
P_T	Total amount of plastic waste that is produced.
PV	Predictor variables used in the MLM.
RF	Random Forest model used to predict MSW_C.
WFM	Waste flow model used to estimate how much of P_E

7.2 Glossary

Following, the most important key words used in this thesis will be briefly explained and clarified.

Municipal solid waste (MSW)	Waste from households, usually including non-hazardous streams of commercial and industrial waste [73]. The composition varies strongly among and also within countries, including organic material, paper, glass, metal, plastic and other compounds. [37]
Sanitary landfill	Government operated facility to where MSW is deposited. At a selected and controlled location, placement is registered, waste is compacted and covered and measures for leachates, final top cover and closure exist. Energy recovery from gases possible [37].
Controlled Landfill	Similar to sanitary landfills but with overall lesser quality. Sites may be less selectively chosen and the control and treatment of leachate and gases might be reduced [37].
Controlled Dump	Non government operated facility where MSW is deposited. This includes site selection, controlled access and sometimes compaction of waste. Neither leachages or gases are monitored or treated. Often includes landfills operated by private agencies [37].
Uncontrolled dumps	Illegal sites where MSW is regularly and directly deposited by both the local population or from within the waste collection system. No engineering measures or controls [37, 73].
Incineration	Burning of MSW to reduce the volume by up to 90%. Controlled incineration controls the emissions and generates energy. Uncontrolled incineration (open burning) creates high air pollution due to the low temperature combustion. Due to the waste composition, addition of fuel might be necessary (high fraction of organic material). Residuals require the use of a sanitary landfill [37].
Recycling	There are two types of recycling: Closed loop recycling, where the waste materials are reused for the same or similar-quality application and cascaded recycling, where the waste products are used in lower-value applications. Recycling rates vary among the waste types: 80-90% for iron and steel, 58% for paper, ~50% for PET plastics (whereas 7% are used in a closed loop recycling), 14% for plastic packaging but only 5% for all plastics combined [80].
Composting	Aerobic process to treat organic waste compounds. Preferably it avoids forming methane formed under aerobic conditions (as opposed to landfills and dumps) [37].

Plastic consists of hundreds of materials with different properties. All of them are based on organic material, either on fossil fuel or on renewable bio materials such as cellulose or vegetable oils [81, 82]. As plastics from renewable sources only make up 1% (4Mt) of the total plastic production in 2015 [1] they are neglected in this work. The most common plastics produced and their use, share of the demand in 2016 in the EU28 plus Norway and Switzerland and density is as follows:

	Type	**Usage** [80, 82]	**Share** [80, 82]	**Density (g/ml)** [74]
Plastic	PET	Drink bottles and food containers	7.4%	1.38-1.41
	PE-HD and PE-MD	Milk bottles, freezer bags	12.3%	0.94-0.97
	PE-LD and PE-LLD	Cling wrap, rubbish bags, reusable bags	17.5%	0.89-0.93
	PVC	Cosmetic containers, pipes, window frames	10.0%	1.16-1.41
	PP	Microwave dishes, sweet and snack wrappers, automotive parts	19.3%	0.85-0.92
	PS and PS-E	Plastic cutlery, egg trays (PS) and foamed polystyrene used for take-away food containers and protective packaging (PS-E)	6.7%	1.04-1.08
	Others	water cooler bottles, flexible films, optical fiber.	19.3%	

7.3 Scatterplots of the Predictor Variables versus MSW_C

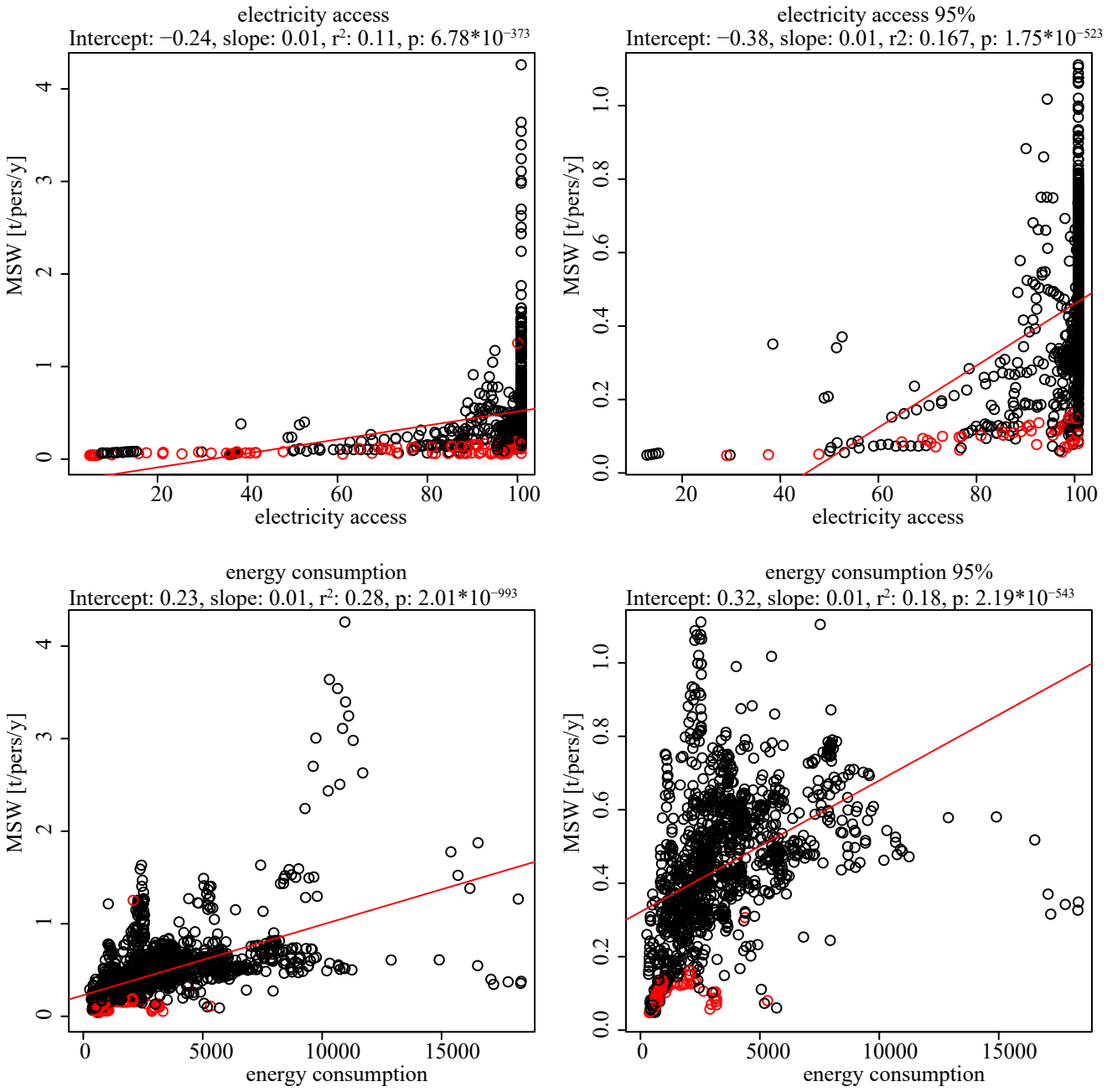

Figure A.1. Continued on next page.

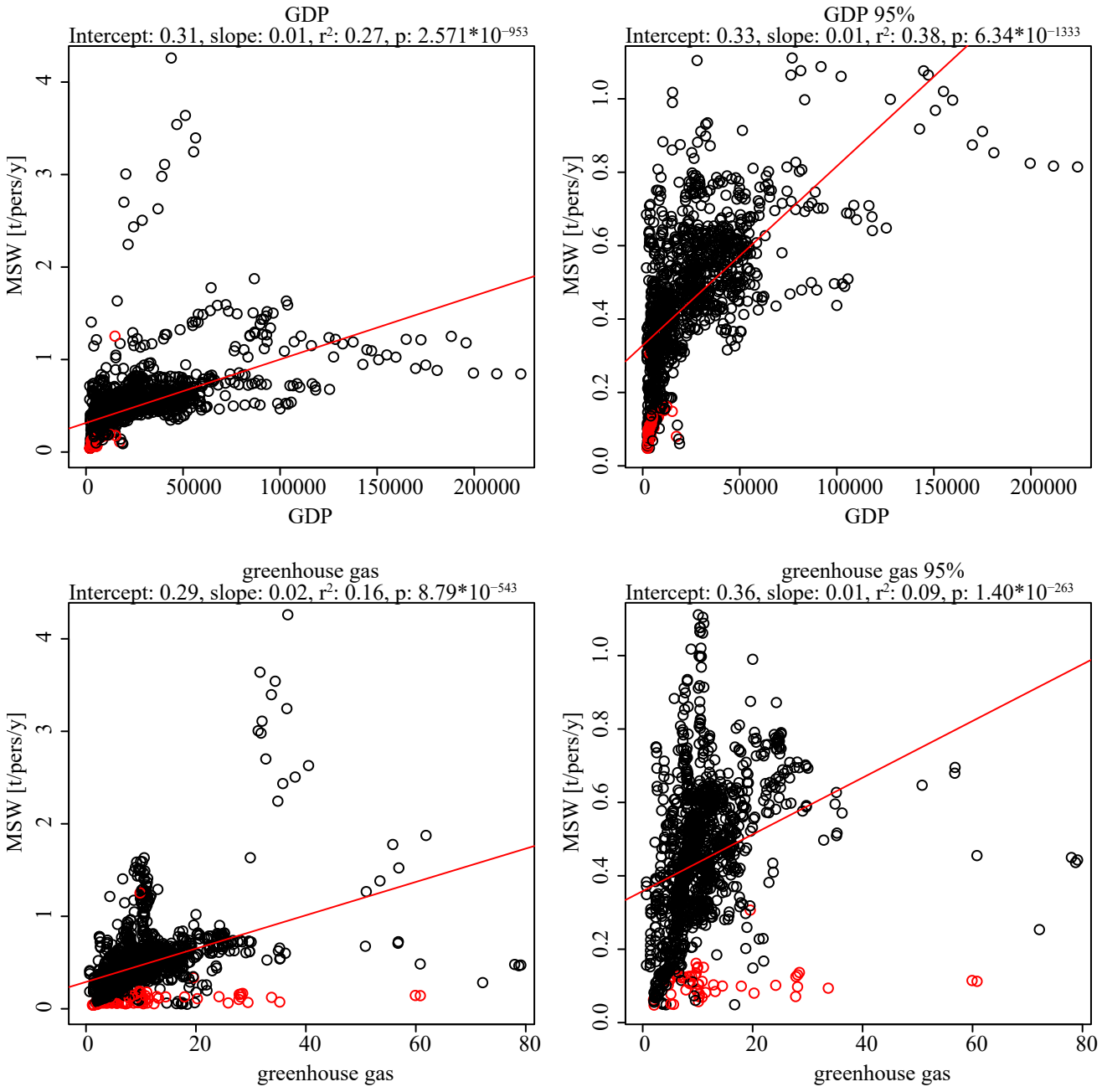

Figure A.1. Continued on next page.

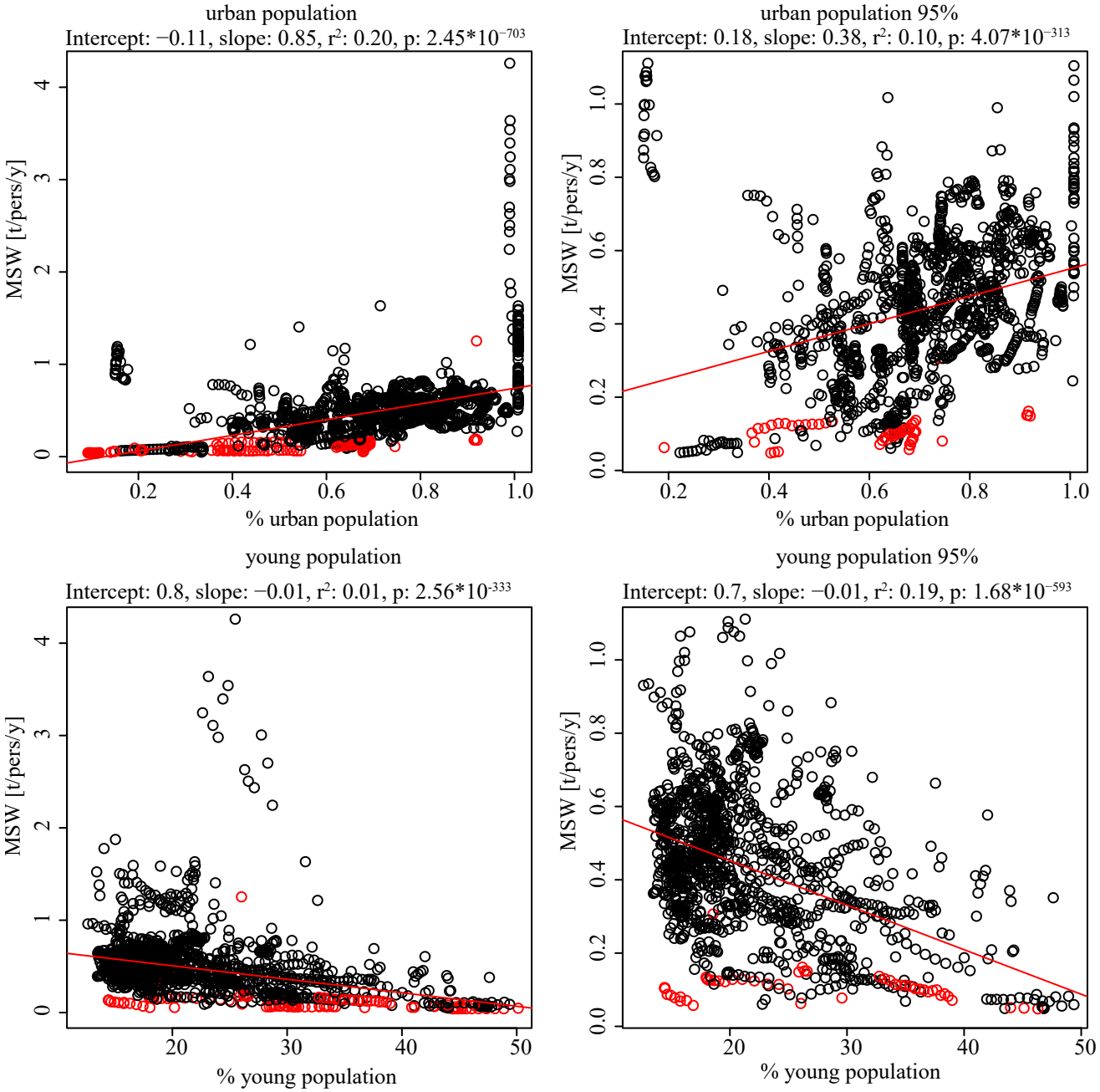

Figure A.1. Continued on next page.

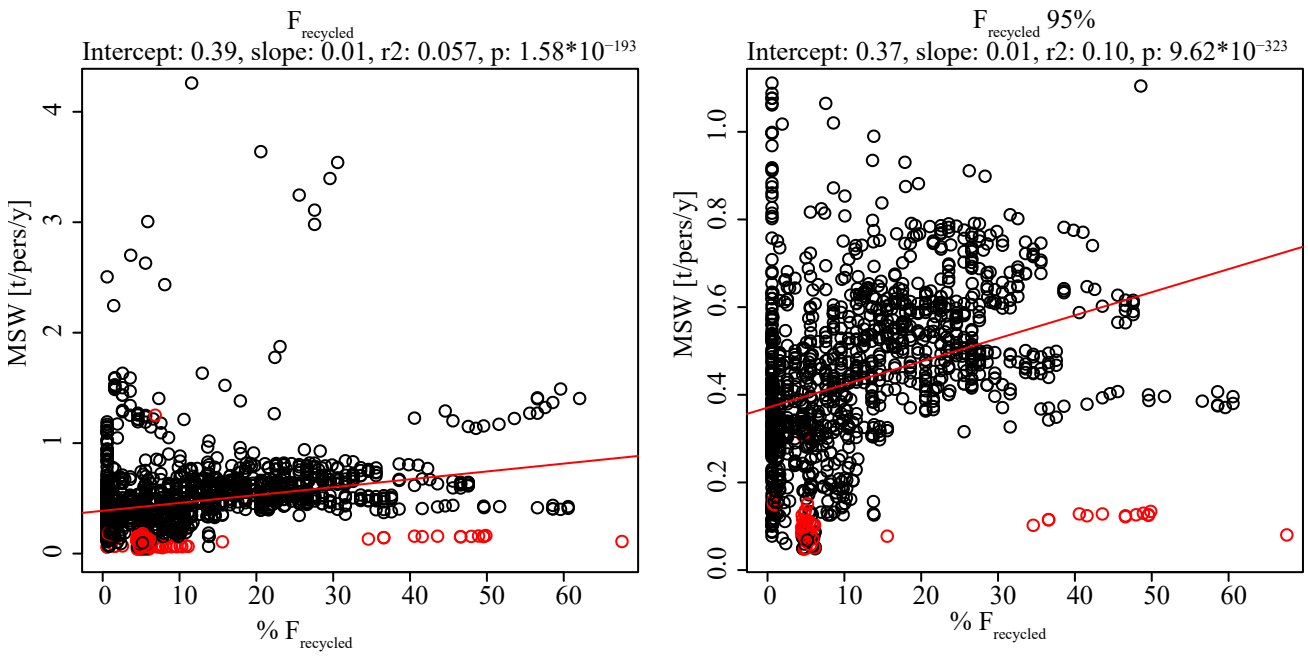

Figure A.1. Scatterplots of MSW_C versus the PVs. The left column includes all 1'394 available data points while the right column includes only the 95% quantile of MSW_C. The red line represents a linear fit and the corresponding r^2 and p-values are indicated in the subtitle of each scatterplot. The red data points represent countries whose prediction error of the RF model is >30%.

7.4 Scatterplots Among the Predictor Variables

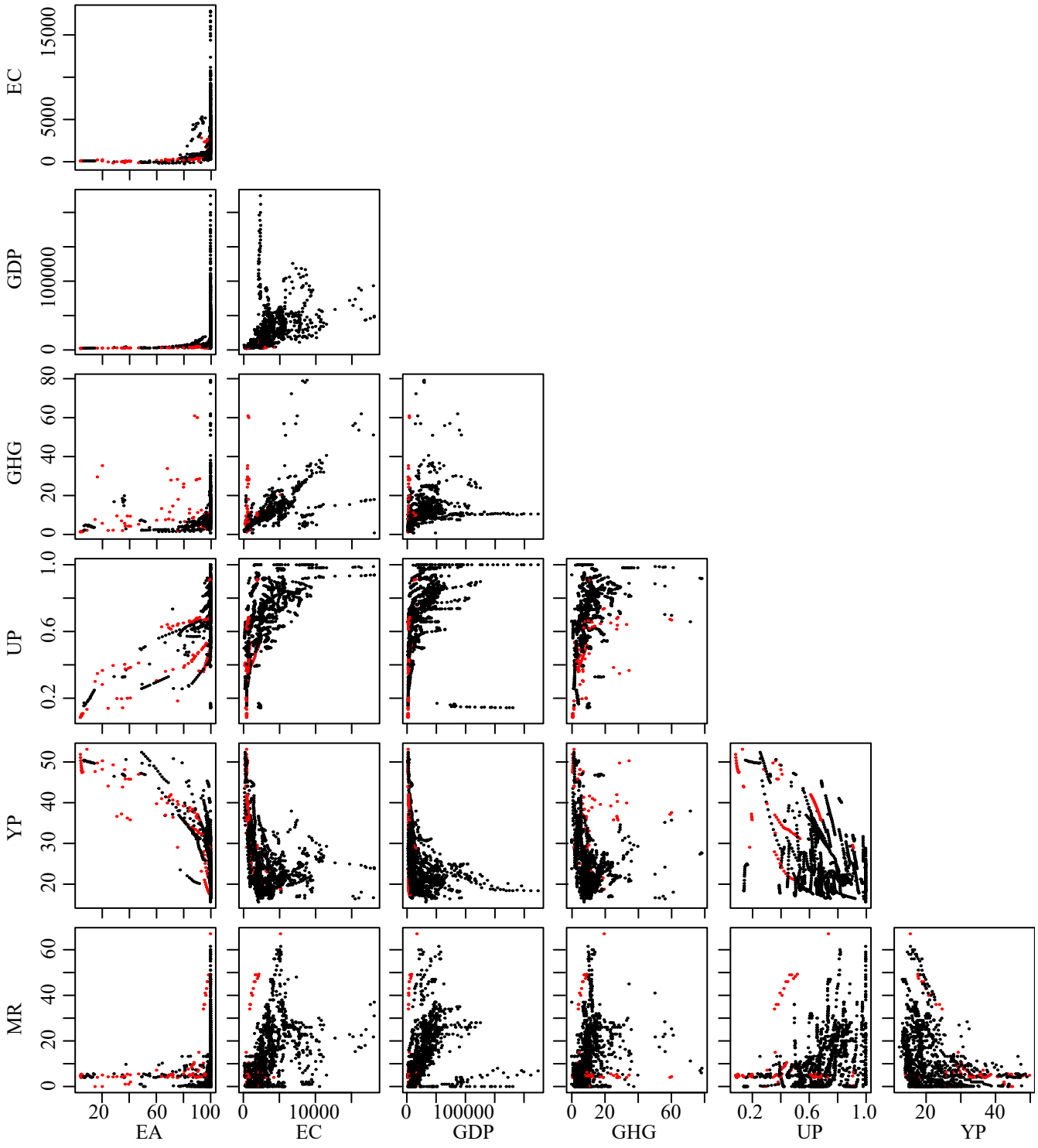

Figure A.2. Scatterplots of the PVs. The red data points represent countries whose prediction error of the RF model is >30%. See figure A.1 for scatterplots of the PVs versus the target variable MSW_C.

7.5 Correlation Dendrogram of the PVs and MSW_C

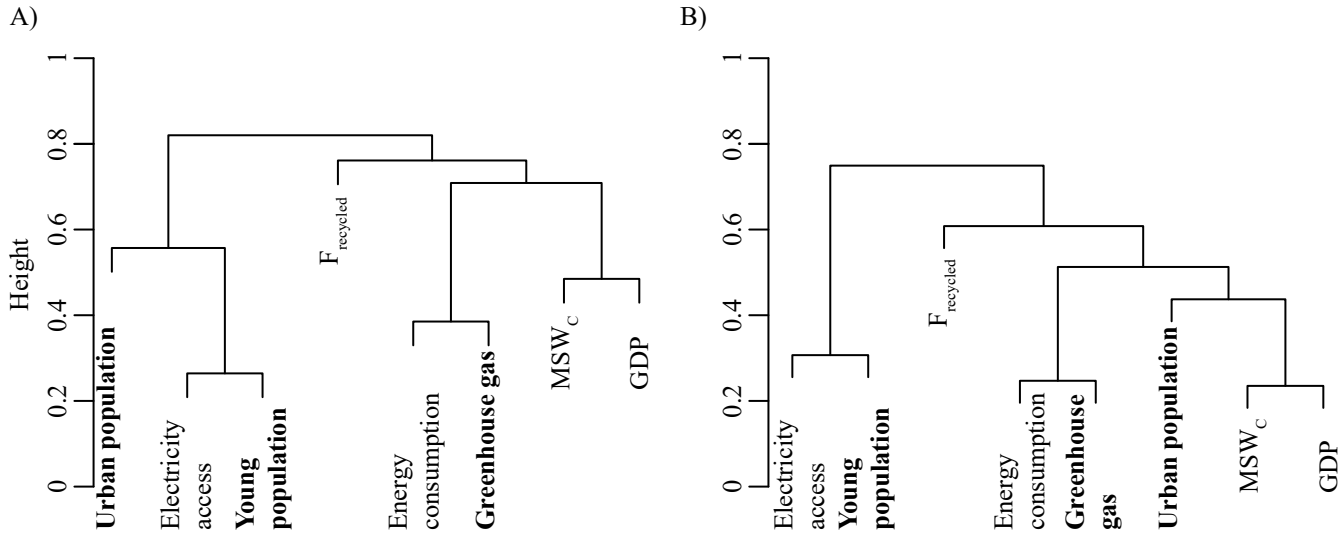

Figure A.3. Dendrogram of the correlations among the target variable MSW_C and the PVs. A) Dendrogram using the Pearson correlations, B) Dendrogram using the Spearman correlations. The PVs in bold font were selected for the MLM as described in section 2.4.2.

7.6 Numbers of GDP-clusters

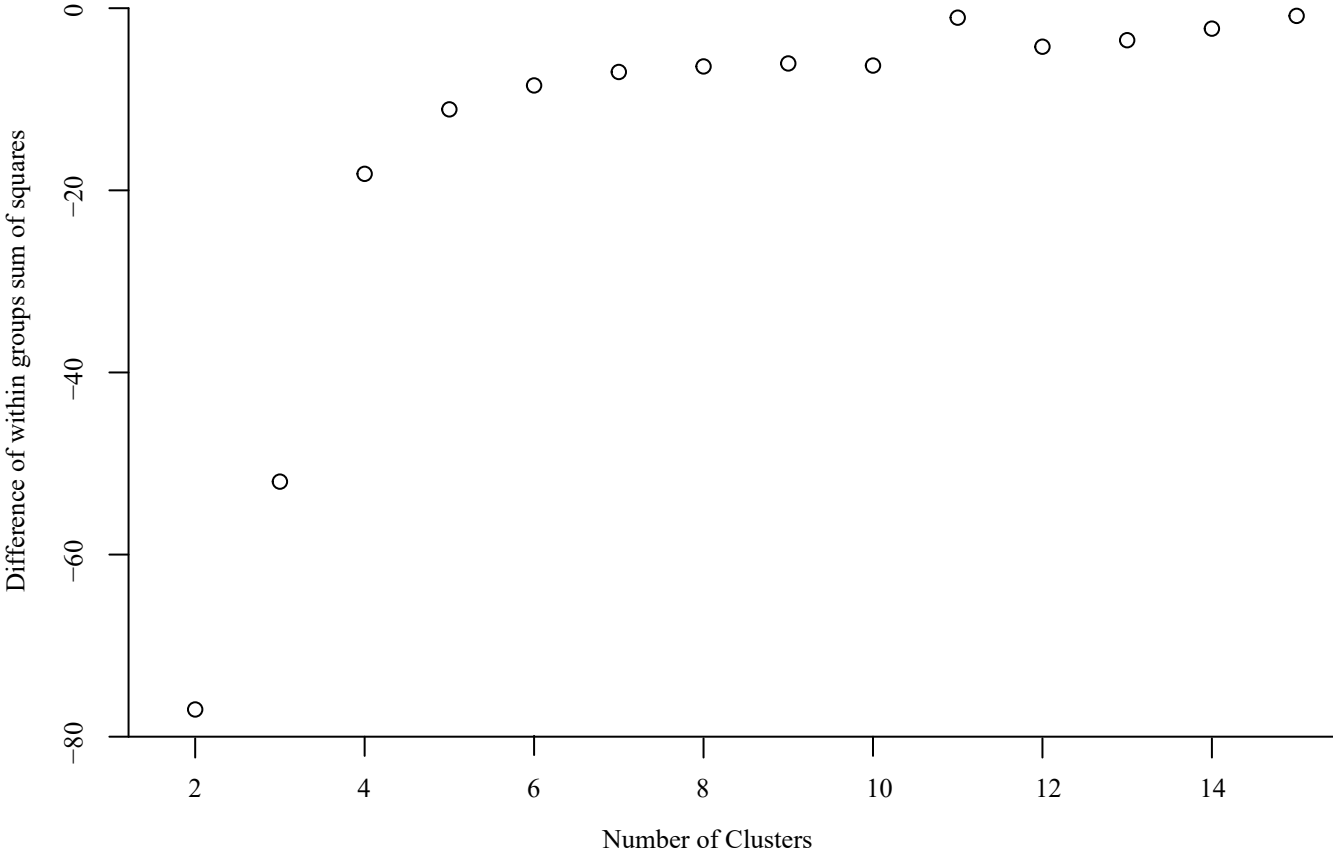

Figure A.4. Differences between the sum of squared errors of the trajectory of each country and the respective cluster trajectory mean for different numbers of clusters (2-15). This was used to determine the number of GDP-clusters as described in section 2.3.3. These clusters are then used to estimate missing data in the PVs of countries that have no data points available.

7.7 NNET Architecture Optimization

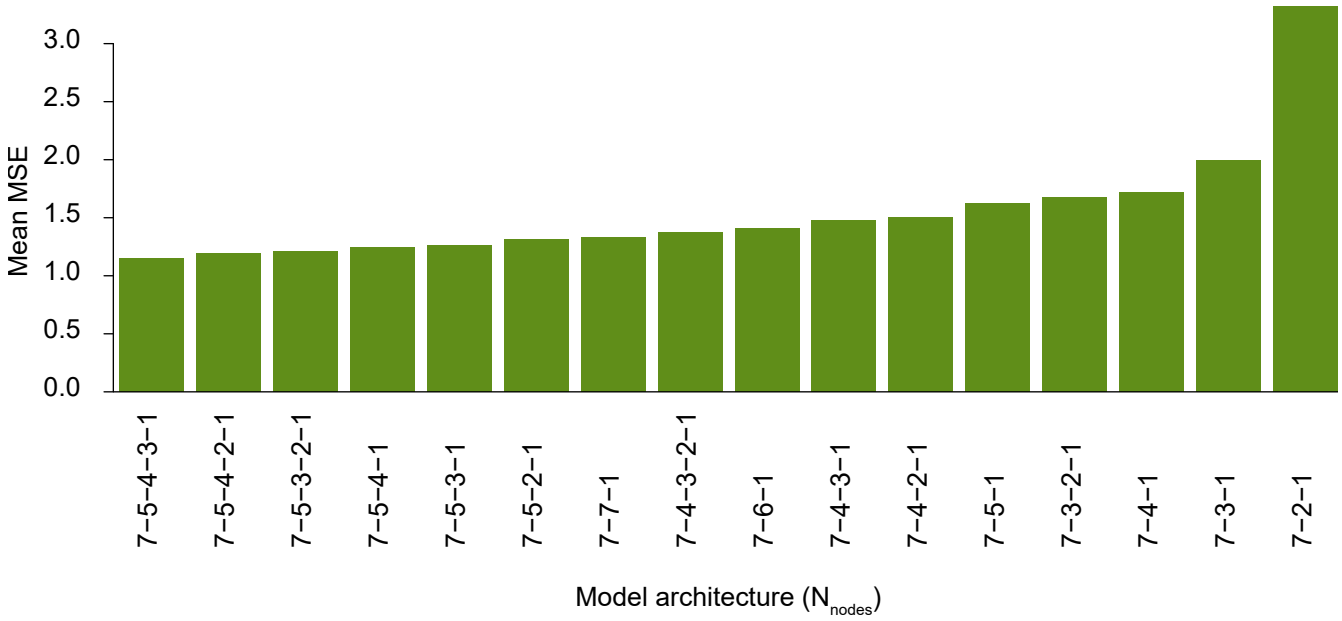

Figure A.5. Summary of the average MSE for the all different layer and node combinations that were tested as possible model architectures for the NNET. The first number indicates the amount of predictor variables, the second number shows the number of nodes in the hidden layer and the third number stands for the number of outputs.

7.8 Number of Trees in RF versus MSE

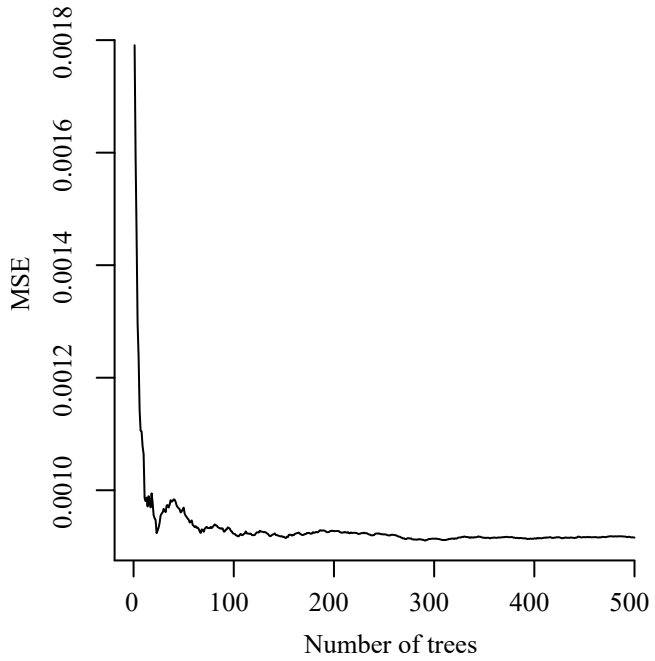

Figure A.6. MSE of the RF versus the number of trees. We chose a number of 500 as the MSE would not decrease by increasing the tree number. This is used in the RF optimization process described in section 2.4.1.

7.9 Data Modification

Table A.1. Overview of the data completeness of each variable and the measures taken to modify the data. If a country has missing data for some but not all years, the data of this variable was inter- and/or extrapolated using either a polynomial (GDP) or linear fit (all others). We indicate the corresponding r_2 and p-vlaues as both mean and median. If the country lacked data in a predictor variable for all years, the average of its corresponding GDP-cluster was applied. The column *N complete* indicates how many of the 173 countries have data available for all years in the time period 1990-2015. The predictor variables $F_{landfills}$, F_{dums} and $F_{plastic}$ do only contain values for the year 2012. Abbreviations: N complete: number of countries with 100% available data data from 1990-2015, N cluster: number of countries where the data was completed using the corresponding GDP-cluster averages.

Predictor variable	**N complete**	**Inter- or extrapolation of data**				**N cluster**
		N	Type	r^2 (mean/median)	p (mean/median)	
Total population	173	0				0
Urban population	173	0				0
Young population	164	1	Lin	0.76 / 0.76	$1.5 \cdot 10^{-7}$ / $1.5 \cdot 10^{-7}$	8
Electricity access	0	170	Lin	0.92 / 0.96	$1.8 \cdot 10^{-4}$ / $4.8 \cdot 10^{-17}$	3
GDP	144	29	Poly	0.90 / 0.92	$4 \cdot 10^{-5}$ / $1.6 \cdot 10^{-13}$	0
GNI	173	0				0
Energy consumption	34	112	Lin	0.57 / 0.62	$9.5 \cdot 10^{-2}$ / $1.4 \cdot 10^{-4}$	27
Greenhouse gas	0	162	Lin	0.42 / 0.39	$1.1 \cdot 10^{-1}$ / $1.6 \cdot 10^{-3}$	11
$F_{collection}$	0	78	Lin	0.70 / 0.74	$8.4 \cdot 10^{-2}$ / $2.2 \cdot 10^{-3}$	95
$F_{recycled}$	0	60	Lin	0.68 / 0.74	$8.8 \cdot 10^{-2}$ / $4.6 \cdot 10^{-6}$	113
$F_{landfills}$, F_{dumps}	68 (2012 only)					105*
$F_{plastic}$	95 (2012 only)					78

** Missing data was not derived from cluster averages but from the corresponding values from [37] based on the GNI.*

7.10 Data Completeness

Table A.2. Overview of the data completeness of each dataset for each country. The years in the column header indicate the period of data availability. Each cell shows haw many data points were available in the whole period (number before the brackets) as well as in the selected period 1990-2015 (in the brackets). Countries with a bold formatting were grouped and those with a cursive formatting were dropped (compare table 4). Note that the variable *energy access* was not available for the year 2015. The variables *GNI*, *Landfills, Dumps* and $MSW_{plastic}$ are only available for one year and it is thus indicated if the data was available or not. Table titles: Cl. = GDP-cluster, Tot. pop. = total population, Urb. pop. = urban population, Y. pop. = young population, E. acc. = electricity access, Collected = $MSW_{collected}$, En. cons. = energy consumption, Gh. gas = greenhouse gas, collection = $MSW_{collection}$, Recyc. = $MSW_{recycled}$, Lf, Dp = Landfills and dumps, Plastic = $MSW_{plastic}$

Country	**Region**	**Cl.**	**Tot. pop. (1950-2050)** [50]	**Urb. pop. (1950-2050)** [51]	**Y. pop. (1960-2016)** [52]	**El. acc. (1990-2014)** [53]	**Collected (1990-2015)** [49]	**GDP (1960-2015)** [54]	**GNI (1960-2016)** [58]	**En. cons. (1990-2015)** [55]	**Gh. gas (1970-2012)** [56]	**Collection (1990-2015)** [59]	**Recyc. (1990, 1995-2013)** [57]	**Lf, Dp (2012)** [37]	**Plastic (2012)** [37]
Country	Continent	Cluster	tot pop (1950-2050)	urb_pop(1950-2050)	young pop(1960-2016)	elec acc (1960-2014)	MSW collected	GDP (1960-2016)	GNI	energy consumption (1990-2015)	GHG (1970-2012)	MSW collection	MSW recycled	Dumps	Plastic
Burundi	Africa	1	101 (26)	101 (26)	57 (26)	25 (25)	11 (11)	57 (26)	L	0 (0)	43 (23)	0 (0)	0 (0)	No data	0
Comoros	Africa	1	101 (26)	101 (26)	57 (26)	25 (25)	0 (0)	37 (26)	L	5 (5)	43 (23)	0 (0)	0 (0)	No data	0
Djibouti	Africa	1	101 (26)	101 (26)	57 (26)	25 (25)	0 (0)	30 (26)	LM	5 (5)	43 (23)	0 (0)	0 (0)	No data	0
Eritrea	Africa	1	101 (26)	101 (26)	52 (22)	25 (25)	0 (0)	20 (20)	L	20 (20)	43 (23)	0 (0)	0 (0)	No data	0
Ethiopia	Africa	1	101 (26)	101 (26)	57 (26)	24 (24)	0 (0)	36 (26)	L	25 (25)	43 (23)	0 (0)	0 (0)	No data	2
Kenya	Africa	1	101 (26)	101 (26)	57 (26)	25 (25)	0 (0)	57 (26)	LM	25 (25)	43 (23)	5 (5)	0 (0)	No data	0
Madagascar	Africa	1	101 (26)	101 (26)	57 (26)	25 (25)	2 (2)	57 (26)	L	0 (0)	43 (23)	3 (3)	3 (3)	0 / 97	1
Malawi	Africa	1	101 (26)	101 (26)	57 (26)	25 (25)	0 (0)	57 (26)	L	0 (0)	43 (23)	0 (0)	0 (0)	No data	0
Mauritius	Africa	3	101 (26)	101 (26)	57 (26)	25 (25)	13 (13)	41 (26)	UM	25 (25)	24 (4)	6 (6)	3 (3)	0 / 91	9
Mayotte	Africa	No data	101 (26)	101 (26)	0 (0)	0 (0)	0 (0)	0 (0)	No data	0 (0)	0 (0)	0 (0)	0 (0)	No data	No data
Mozambique	Africa	1	101 (26)	101 (26)	57 (26)	24 (24)	0 (0)	37 (26)	L	25 (25)	43 (23)	0 (0)	0 (0)	No data	10

Table A.2 – Continued from previous page

Country	Region	Cl.	Tot. pop.	Urb. pop.	Y. pop.	El. acc.	Collec-ted	GDP	GNI	En. cons.	Gh. gas	Collec-tion	Recyc.	Lf, Dumps	Plastic
Réunion	Africa	No data	101 (26)	101 (26)	0 (0)	0 (0)	8 (8)	0 (0)	No data	0 (0)	0 (0)	4 (4)	0 (0)	No data	No data
Rwanda	Africa	1	101 (26)	101 (26)	57 (26)	24 (24)	0 (0)	57 (26)	L	0 (0)	43 (23)	0 (0)	0 (0)	No data	0
Seychelles	Africa	3	101 (26)	101 (26)	57 (26)	25 (25)	0 (0)	57 (26)	H	5 (5)	36 (16)	0 (0)	0 (0)	No data	0
Somalia	Africa	1	101 (26)	101 (26)	57 (26)	23 (23)	0 (0)	4 (3)	L	0 (0)	43 (23)	0 (0)	0 (0)	No data	0
South Sudan	Africa	1	101 (26)	101 (26)	57 (26)	13 (13)	0 (0)	8 (8)	L	3 (3)	0 (0)	0 (0)	0 (0)	No data	0
Uganda	Africa	1	101 (26)	101 (26)	57 (26)	25 (25)	1 (1)	29 (26)	L	25 (25)	43 (23)	0 (0)	0 (0)	0 / 100	1
United Republic of Tanzania	Africa	1	101 (26)	101 (26)	0 (0)	25 (25)	0 (0)	27 (26)	L	0 (0)	43 (23)	0 (0)	0 (0)	No data	No data
Zambia	Africa	1	101 (26)	101 (26)	57 (26)	25 (25)	2 (2)	57 (26)	LM	25 (25)	32 (18)	4 (4)	0 (0)	No data	5
Zimbabwe	Africa	1	101 (26)	101 (26)	57 (26)	25 (25)	0 (0)	57 (26)	L	24 (24)	43 (23)	0 (0)	0 (0)	No data	20
Angola	Africa	2	101 (26)	101 (26)	57 (26)	25 (25)	1 (1)	32 (26)	UM	25 (25)	43 (23)	1 (1)	0 (0)	No data	0
Cameroon	Africa	1	101 (26)	101 (26)	57 (26)	25 (25)	4 (4)	57 (26)	LM	25 (25)	43 (23)	2 (2)	4 (4)	95 / 0	5
Central African Republic	Africa	1	101 (26)	101 (26)	57 (26)	25 (25)	0 (0)	57 (26)	L	0 (0)	43 (23)	0 (0)	0 (0)	No data	No data
Chad	Africa	1	101 (26)	101 (26)	57 (26)	24 (24)	0 (0)	57 (26)	L	0 (0)	43 (23)	0 (0)	0 (0)	No data	0
Congo	Africa	1	101 (26)	101 (26)	57 (26)	25 (25)	0 (0)	57 (26)	LM	25 (25)	43 (23)	0 (0)	0 (0)	No data	No data
Democratic Republic of the Congo	Africa	1	101 (26)	101 (26)	57 (26)	24 (24)	0 (0)	56 (25)	L	25 (25)	43 (23)	0 (0)	0 (0)	No data	No data
Equatorial Guinea	Africa	5	101 (26)	101 (26)	57 (26)	25 (25)	0 (0)	53 (26)	UM	4 (4)	40 (20)	0 (0)	0 (0)	No data	0
Gabon	Africa	3	101 (26)	101 (26)	57 (26)	25 (25)	0 (0)	57 (26)	UM	25 (25)	43 (23)	0 (0)	0 (0)	No data	0
Sao Tome and Principe	Africa	1	101 (26)	101 (26)	57 (26)	25 (25)	0 (0)	16 (15)	No data	5 (5)	43 (23)	0 (0)	0 (0)	No data	0

Table A.2 – Continued from previous page

Country	Region	Cl.	Tot. pop.	Urb. pop.	Y. pop.	El. acc.	Collec-ted	GDP	GNI	En. cons.	Gh. gas	Collec-tion	Recyc.	Lf, Dumps	Plastic
Algeria	Africa	2	101 (26)	101 (26)	57 (26)	25 (25)	3 (3)	57 (26)	UM	25 (25)	43 (23)	4 (4)	1 (1)	97 / 0	5
Egypt	Africa	2	101 (26)	101 (26)	57 (26)	25 (25)	4 (4)	52 (26)	LM	25 (25)	43 (23)	0 (0)	3 (3)	No data	12
Libya	Africa	3	101 (26)	101 (26)	57 (26)	25 (25)	0 (0)	22 (22)	UM	25 (25)	43 (23)	0 (0)	0 (0)	No data	0
Morocco	Africa	1	101 (26)	101 (26)	57 (26)	25 (25)	3 (3)	57 (26)	LM	25 (25)	43 (23)	0 (0)	1 (1)	95 / 1	4
Sudan	Africa	1	101 (26)	101 (26)	57 (26)	25 (25)	8 (8)	57 (26)	LM	25 (25)	43 (23)	0 (0)	0 (0)	No data	0
Tunisia	Africa	2	101 (26)	101 (26)	57 (26)	25 (25)	6 (6)	52 (26)	LM	25 (25)	43 (23)	4 (4)	2 (2)	58 / 42	11
Western Sahara	Africa	No data	101 (26)	101 (26)	0 (0)	0 (0)	0 (0)	0 (0)	No data	0 (0)	0 (0)	0 (0)	0 (0)	No data	No data
Botswana	Africa	2	101 (26)	101 (26)	57 (26)	25 (25)	0 (0)	57 (26)	UM	25 (25)	43 (23)	0 (0)	0 (0)	No data	0
Lesotho	Africa	1	101 (26)	101 (26)	57 (26)	17 (17)	0 (0)	57 (26)	LM	4 (4)	43 (23)	1 (1)	0 (0)	No data	0
Namibia	Africa	2	101 (26)	101 (26)	57 (26)	25 (25)	0 (0)	37 (26)	UM	24 (24)	43 (23)	0 (0)	0 (0)	No data	0
South Africa	Africa	2	101 (26)	101 (26)	57 (26)	25 (25)	0 (0)	27 (26)	UM	24 (24)	43 (23)	0 (0)	0 (0)	No data	0
Swaziland	Africa	2	101 (26)	101 (26)	57 (26)	24 (24)	0 (0)	57 (26)	LM	5 (5)	43 (23)	0 (0)	0 (0)	No data	0
Benin	Africa	1	101 (26)	101 (26)	57 (26)	25 (25)	0 (0)	57 (26)	L	25 (25)	43 (23)	0 (0)	0 (0)	No data	7
Burkina Faso	Africa	1	101 (26)	101 (26)	57 (26)	25 (25)	14 (14)	57 (26)	L	0 (0)	43 (23)	6 (6)	0 (0)	No data	0
Cabo Verde	Africa	2	101 (26)	101 (26)	57 (26)	25 (25)	11 (11)	37 (26)	LM	5 (5)	38 (18)	11 (11)	0 (0)	No data	0
Côte d'Ivoire	Africa	1	101 (26)	101 (26)	57 (26)	25 (25)	0 (0)	57 (26)	LM	25 (25)	43 (23)	0 (0)	0 (0)	No data	No data
Gambia	Africa	1	101 (26)	101 (26)	57 (26)	25 (25)	0 (0)	51 (26)	L	5 (5)	43 (23)	0 (0)	0 (0)	No data	0
Ghana	Africa	1	101 (26)	101 (26)	57 (26)	25 (25)	0 (0)	57 (26)	LM	25 (25)	43 (23)	0 (0)	0 (0)	No data	4
Guinea	Africa	1	101 (26)	101 (26)	57 (26)	25 (25)	0 (0)	31 (26)	L	0 (0)	43 (23)	0 (0)	0 (0)	No data	4
Guinea-Bissau	Africa	1	101 (26)	101 (26)	57 (26)	11 (11)	0 (0)	47 (26)	L	5 (5)	43 (23)	0 (0)	0 (0)	No data	0

Table A.2 – Continued from previous page

Country	Region	Cl.	Tot. pop.	Urb. pop.	Y. pop.	El. acc.	Collec-ted	GDP	GNI	En. cons.	Gh. gas	Collec-tion	Recyc.	Lf, Dumps	Plastic
Liberia	Africa	1	101 (26)	101 (26)	57 (26)	12 (12)	0 (0)	57 (26)	L	0 (0)	43 (23)	0 (0)	0 (0)	No data	13
Mali	Africa	1	101 (26)	101 (26)	57 (26)	24 (24)	0 (0)	50 (26)	L	0 (0)	43 (23)	0 (0)	0 (0)	No data	2
Mauritania	Africa	1	101 (26)	101 (26)	57 (26)	25 (25)	0 (0)	57 (26)	LM	0 (0)	43 (23)	0 (0)	0 (0)	No data	0
Niger	Africa	No data	101 (26)	101 (26)	57 (26)	25 (25)	8 (8)	0 (0)	L	15 (15)	43 (23)	0 (0)	8 (8)	0 / 64	2
Nigeria	Africa	1	101 (26)	101 (26)	57 (26)	25 (25)	0 (0)	57 (26)	LM	25 (25)	43 (23)	0 (0)	0 (0)	No data	18
Saint Helena	Africa	No data	101 (26)	101 (26)	0 (0)	0 (0)	0 (0)	0 (0)	No data	0 (0)	0 (0)	0 (0)	0 (0)	No data	No data
Senegal	Africa	1	101 (26)	101 (26)	57 (26)	25 (25)	3 (3)	57 (26)	L	25 (25)	43 (23)	8 (8)	0 (0)	No data	3
Sierra Leone	Africa	1	101 (26)	101 (26)	57 (26)	25 (25)	0 (0)	57 (26)	L	0 (0)	43 (23)	0 (0)	0 (0)	No data	0
Togo	Africa	1	101 (26)	101 (26)	57 (26)	24 (24)	3 (3)	57 (26)	L	25 (25)	43 (23)	0 (0)	2 (2)	No data	10
China	Asia	2	101 (26)	101 (26)	57 (26)	25 (25)	13 (13)	57 (26)	UM	25 (25)	43 (23)	0 (0)	0 (0)	No data	0
China, Hong Kong SAR	Asia	4	101 (26)	101 (26)	57 (26)	25 (25)	16 (16)	57 (26)	H	25 (25)	43 (23)	16 (16)	16 (16)	No data	No data
China, Macao SAR	Asia	6	101 (26)	101 (26)	57 (26)	21 (21)	16 (16)	35 (26)	H	0 (0)	14 (2)	16 (16)	30 (30)	No data	No data
Dem. People's Republic of Korea	Asia	No data	101 (26)	101 (26)	57 (26)	25 (25)	0 (0)	0 (0)	L	25 (25)	43 (23)	0 (0)	0 (0)	No data	No data
Japan	Asia	4	101 (26)	101 (26)	57 (26)	25 (25)	24 (24)	57 (26)	H	26 (26)	43 (23)	2 (2)	20 (20)	0 / 3	9
Mongolia	Asia	2	101 (26)	101 (26)	57 (26)	25 (25)	0 (0)	36 (26)	LM	25 (25)	41 (21)	0 (0)	0 (0)	No data	0
Republic of Korea	Asia	5	101 (26)	101 (26)	57 (26)	20 (20)	24 (24)	57 (26)	H	26 (26)	43 (23)	1 (1)	18 (18)	No data	No data
Kazakhstan	Asia	3	101 (26)	101 (26)	57 (26)	25 (25)	5 (5)	27 (26)	UM	25 (25)	43 (23)	0 (0)	0 (0)	No data	0
Kyrgyzstan	Asia	No data	101 (26)	101 (26)	57 (26)	25 (25)	19 (19)	0 (0)	LM	25 (25)	43 (23)	0 (0)	0 (0)	No data	No data

Table A.2 – Continued from previous page

Country	Region	Cl.	Tot. pop.	Urb. pop.	Y. pop.	El. acc.	Collec-ted	GDP	GNI	En. cons.	Gh. gas	Collec-tion	Recyc.	Lf, Dumps	Plastic
Tajikistan	Asia	1	101 (26)	101 (26)	57 (26)	25 (25)	0 (0)	27 (26)	LM	25 (25)	43 (23)	0 (0)	0 (0)	No data	0
Turkmenistan	Asia	2	101 (26)	101 (26)	57 (26)	25 (25)	0 (0)	30 (26)	UM	25 (25)	43 (23)	0 (0)	0 (0)	No data	0
Uzbekistan	Asia	1	101 (26)	101 (26)	57 (26)	25 (25)	0 (0)	57 (26)	LM	26 (26)	43 (23)	0 (0)	0 (0)	No data	0
Afghanistan	Asia	1	101 (26)	101 (26)	57 (26)	16 (16)	0 (0)	38 (15)	L	0 (0)	42 (22)	0 (0)	0 (0)	No data	0
Bangladesh	Asia	1	101 (26)	101 (26)	57 (26)	25 (25)	0 (0)	57 (26)	LM	25 (25)	43 (23)	0 (0)	0 (0)	No data	7
Bhutan	Asia	1	101 (26)	101 (26)	57 (26)	24 (24)	1 (1)	37 (26)	LM	5 (5)	43 (23)	1 (1)	1 (1)	No data	13
India	Asia	1	101 (26)	101 (26)	57 (26)	25 (25)	0 (0)	57 (26)	LM	25 (25)	43 (23)	0 (0)	0 (0)	No data	2
Iran (Islamic Republic of)	Asia	2	101 (26)	101 (26)	57 (26)	25 (25)	0 (0)	55 (24)	UM	25 (25)	38 (18)	0 (0)	0 (0)	No data	No data
Maldives	Asia	3	101 (26)	101 (26)	57 (26)	25 (25)	16 (16)	37 (26)	UM	5 (5)	20 (5)	5 (5)	0 (0)	No data	0
Nepal	Asia	1	101 (26)	101 (26)	57 (26)	24 (24)	2 (2)	57 (26)	L	25 (25)	43 (23)	0 (0)	0 (0)	No data	3
Pakistan	Asia	1	101 (26)	101 (26)	57 (26)	25 (25)	0 (0)	57 (26)	LM	25 (25)	43 (23)	0 (0)	0 (0)	No data	18
Sri Lanka	Asia	2	101 (26)	101 (26)	57 (26)	25 (25)	1 (1)	57 (26)	LM	25 (25)	43 (23)	0 (0)	0 (0)	No data	6
Brunei Darussalam	Asia	4	101 (26)	101 (26)	57 (26)	25 (25)	7 (7)	52 (26)	H	25 (25)	43 (23)	0 (0)	0 (0)	No data	2
Cambodia	Asia	1	101 (26)	101 (26)	57 (26)	24 (24)	4 (4)	39 (23)	LM	20 (20)	43 (23)	0 (0)	0 (0)	100 / 0	10
Indonesia	Asia	2	101 (26)	101 (26)	57 (26)	25 (25)	21 (21)	50 (26)	LM	25 (25)	43 (23)	0 (0)	0 (0)	No data	10
Lao People's Democratic Republic	Asia	1	101 (26)	101 (26)	57 (26)	25 (25)	0 (0)	33 (26)	LM	0 (0)	43 (23)	0 (0)	0 (0)	No data	No data
Malaysia	Asia	3	101 (26)	101 (26)	57 (26)	25 (25)	16 (16)	57 (26)	UM	25 (25)	43 (23)	0 (0)	0 (0)	No data	12
Myanmar	Asia	1	101 (26)	101 (26)	57 (26)	25 (25)	0 (0)	17 (16)	LM	25 (25)	43 (23)	0 (0)	0 (0)	No data	16

Table A.2 – Continued from previous page

Country	Region	Cl.	Tot. pop.	Urb. pop.	Y. pop.	El. acc.	Collec-ted	GDP	GNI	En. cons.	Gh. gas	Collec-tion	Recyc.	Lf, Dumps	Plastic
Philippines	Asia	1	101 (26)	101 (26)	57 (26)	25 (25)	3 (3)	57 (26)	LM	25 (25)	43 (23)	3 (3)	0 (0)	No data	14
Singapore	Asia	4	101 (26)	101 (26)	57 (26)	25 (25)	20 (20)	57 (26)	H	25 (25)	11 (11)	20 (20)	19 (19)	0 / 15	12
Thailand	Asia	2	101 (26)	101 (26)	57 (26)	25 (25)	0 (0)	57 (26)	UM	25 (25)	43 (23)	0 (0)	0 (0)	33 / 59	14
Timor-Leste	Asia	1	101 (26)	101 (26)	57 (26)	25 (25)	0 (0)	17 (16)	LM	4 (4)	43 (23)	0 (0)	0 (0)	No data	0
Viet Nam	Asia	1	101 (26)	101 (26)	57 (26)	25 (25)	0 (0)	48 (18)	LM	0 (0)	43 (23)	0 (0)	0 (0)	No data	No data
Armenia	Asia	2	101 (26)	101 (26)	57 (26)	25 (25)	19 (19)	27 (26)	LM	25 (25)	43 (23)	19 (19)	16 (16)	0 / 100	10
Azerbaijan	Asia	2	101 (26)	101 (26)	57 (26)	25 (25)	13 (13)	27 (26)	UM	25 (25)	42 (22)	0 (0)	0 (0)	No data	0
Bahrain	Asia	5	101 (26)	101 (26)	57 (26)	14 (14)	0 (0)	37 (26)	H	25 (25)	38 (18)	0 (0)	0 (0)	No data	0
Cyprus	Asia	5	101 (26)	101 (26)	57 (26)	25 (25)	20 (20)	42 (26)	H	25 (25)	41 (21)	14 (14)	19 (19)	0 / 87	11
Georgia	Asia	2	101 (26)	101 (26)	57 (26)	25 (25)	3 (3)	27 (26)	UM	25 (25)	43 (23)	9 (9)	0 (0)	No data	3
Iraq	Asia	3	101 (26)	101 (26)	57 (26)	25 (25)	7 (7)	41 (13)	UM	25 (25)	43 (23)	7 (7)	0 (0)	No data	0
Israel	Asia	4	101 (26)	101 (26)	57 (26)	25 (25)	19 (19)	57 (26)	H	26 (26)	43 (23)	0 (0)	10 (10)	0 / 90	13
Jordan	Asia	2	101 (26)	101 (26)	57 (26)	25 (25)	6 (6)	52 (26)	UM	25 (25)	43 (23)	0 (0)	0 (0)	0 / 85	16
Kuwait	Asia	4	101 (26)	101 (26)	57 (26)	25 (25)	13 (13)	52 (26)	H	22 (22)	41 (21)	19 (19)	18 (18)	No data	0
Lebanon	Asia	3	101 (26)	101 (26)	57 (26)	25 (25)	4 (4)	29 (26)	UM	25 (25)	35 (17)	0 (0)	4 (4)	37 / 46	7
State of Palestine	Asia	No data	101 (26)	101 (26)	0 (0)	0 (0)	15 (15)	0 (0)	No data	0 (0)	0 (0)	11 (11)	24 (24)	No data	No data
Oman	Asia	5	101 (26)	101 (26)	57 (26)	8 (8)	0 (0)	52 (26)	H	25 (25)	43 (23)	0 (0)	0 (0)	No data	0
Qatar	Asia	6	101 (26)	101 (26)	57 (26)	25 (25)	5 (5)	47 (26)	H	25 (25)	37 (17)	5 (5)	2 (2)	No data	0
Saudi Arabia	Asia	5	101 (26)	101 (26)	57 (26)	11 (11)	0 (0)	49 (26)	H	25 (25)	28 (13)	0 (0)	0 (0)	No data	0
Syrian Arab Republic	Asia	2	101 (26)	101 (26)	57 (26)	25 (25)	0 (0)	48 (18)	LM	25 (25)	42 (23)	2 (2)	0 (0)	No data	No data
Turkey	Asia	3	101 (26)	101 (26)	57 (26)	25 (25)	22 (22)	57 (26)	UM	26 (26)	43 (23)	15 (15)	20 (20)	66 / 30	9
United Arab Emirates	Asia	4	101 (26)	101 (26)	57 (26)	25 (25)	6 (6)	42 (26)	H	25 (25)	43 (23)	1 (1)	3 (3)	No data	0

Table A.2 – Continued from previous page

Country	Region	Cl.	Tot. pop.	Urb. pop.	Y. pop.	El. acc.	Collec-ted	GDP	GNI	En. cons.	Gh. gas	Collec-tion	Recyc.	Lf, Dumps	Plastic
West Bank and Gaza	Asia	No data	0 (0)	0 (0)	27 (26)	25 (25)	0 (0)	23 (22)	LM	0 (0)	0 (0)	0 (0)	0 (0)	No data	No data
Yemen	Asia	1	101 (26)	101 (26)	0 (0)	25 (25)	14 (14)	17 (16)	LM	15 (15)	0 (0)	11 (11)	0 (0)	No data	0
Belarus	Europe	2	101 (26)	101 (26)	57 (26)	25 (25)	15 (15)	27 (26)	UM	25 (25)	43 (23)	17 (17)	0 (0)	0 / 96	10
Bulgaria	Europe	2	101 (26)	101 (26)	57 (26)	25 (25)	20 (20)	37 (26)	UM	25 (25)	43 (23)	15 (15)	31 (31)	0 / 83	0
Czech Republic	Europe	5	101 (26)	101 (26)	57 (26)	25 (25)	21 (21)	27 (26)	H	26 (26)	43 (23)	12 (12)	27 (27)	0 / 80	4
Hungary	Europe	3	101 (26)	101 (26)	57 (26)	25 (25)	21 (21)	26 (25)	H	26 (26)	43 (23)	17 (17)	22 (22)	0 / 90	17
Poland	Europe	3	101 (26)	101 (26)	57 (26)	25 (25)	25 (25)	27 (26)	H	26 (26)	43 (23)	7 (7)	27 (27)	0 / 92	10
Republic of Moldova	Europe	1	101 (26)	101 (26)	57 (26)	25 (25)	18 (18)	22 (21)	LM	25 (25)	43 (23)	0 (0)	0 (0)	No data	No data
Romania	Europe	2	101 (26)	101 (26)	57 (26)	25 (25)	20 (20)	30 (26)	UM	25 (25)	43 (23)	14 (14)	29 (29)	0 / 75	3
Russian Federation	Europe	3	101 (26)	101 (26)	57 (26)	25 (25)	2 (2)	28 (26)	UM	25 (25)	43 (23)	0 (0)	4 (4)	No data	No data
Slovakia	Europe	3	101 (26)	101 (26)	57 (26)	25 (25)	22 (22)	27 (26)	H	26 (26)	43 (23)	9 (9)	18 (18)	No data	7
Ukraine	Europe	1	101 (26)	101 (26)	57 (26)	25 (25)	12 (12)	57 (26)	LM	0 (0)	43 (23)	0 (0)	1 (1)	No data	0
Channel Islands	Europe	No data	101 (26)	101 (26)	57 (26)	25 (25)	0 (0)	10 (10)	H	0 (0)	0 (0)	0 (0)	0 (0)	No data	No data
Denmark	Europe	4	101 (26)	101 (26)	57 (26)	25 (25)	21 (21)	57 (26)	H	26 (26)	43 (23)	15 (15)	21 (21)	0 / 5	1
Estonia	Europe	5	101 (26)	101 (26)	57 (26)	25 (25)	20 (20)	22 (21)	H	26 (26)	43 (23)	15 (15)	29 (29)	No data	0
Faroe Islands	Europe	4	101 (26)	101 (26)	0 (0)	0 (0)	0 (0)	18 (18)	No data	0 (0)	0 (0)	0 (0)	0 (0)	No data	0
Finland	Europe	4	101 (26)	101 (26)	57 (26)	25 (25)	21 (21)	57 (26)	H	26 (26)	43 (23)	16 (16)	19 (19)	No data	10
Iceland	Europe	4	101 (26)	101 (26)	57 (26)	25 (25)	19 (19)	57 (26)	H	26 (26)	43 (23)	17 (17)	16 (16)	0 / 72	17
Ireland	Europe	4	101 (26)	101 (26)	57 (26)	25 (25)	18 (18)	57 (26)	H	26 (26)	43 (23)	1 (1)	26 (26)	0 / 66	11
Isle of Man	Europe	6	101 (26)	101 (26)	0 (0)	25 (25)	0 (0)	21 (21)	H	0 (0)	0 (0)	0 (0)	0 (0)	No data	0
Latvia	Europe	3	101 (26)	101 (26)	57 (26)	25 (25)	20 (20)	22 (21)	H	25 (25)	43 (23)	15 (15)	13 (13)	60 / 40	0
Lithuania	Europe	3	101 (26)	101 (26)	57 (26)	25 (25)	20 (20)	22 (21)	H	25 (25)	43 (23)	3 (3)	22 (22)	0 / 44	0
Norway	Europe	6	101 (26)	101 (26)	57 (26)	25 (25)	24 (24)	57 (26)	H	26 (26)	43 (23)	16 (16)	21 (21)	0 / 26	9
Sweden	Europe	4	101 (26)	101 (26)	57 (26)	25 (25)	22 (22)	57 (26)	H	26 (26)	43 (23)	6 (6)	20 (20)	0 / 5	2

Table A.2 – Continued from previous page

Country	Region	Cl.	Tot. pop.	Urb. pop.	Y. pop.	El. acc.	Collec-ted	GDP	GNI	En. cons.	Gh. gas	Collec-tion	Recyc.	Lf, Dumps	Plastic
United King-dom	Europe	4	101 (26)	101 (26)	57 (26)	25 (25)	21 (21)	57 (26)	H	26 (26)	43 (23)	14 (14)	21 (21)	0 / 64	0
Albania	Europe	2	101 (26)	101 (26)	57 (26)	25 (25)	10 (10)	33 (26)	UM	25 (25)	43 (23)	5 (5)	0 (0)	No data	8
Andorra	Europe	4	101 (26)	101 (26)	0 (0)	25 (25)	18 (18)	47 (26)	H	0 (0)	0 (0)	22 (22)	18 (18)	No data	14
Bosnia and Herzegovina	Europe	2	101 (26)	101 (26)	57 (26)	25 (25)	8 (8)	23 (22)	UM	25 (25)	43 (23)	8 (8)	0 (0)	No data	0
Croatia	Europe	3	101 (26)	101 (26)	57 (26)	25 (25)	14 (14)	22 (21)	H	25 (25)	43 (23)	9 (9)	9 (9)	0 / 70	12
Gibraltar	Europe	No data	101 (26)	101 (26)	0 (0)	25 (25)	0 (0)	0 (0)	H	25 (25)	29 (18)	0 (0)	0 (0)	No data	0
Greece	Europe	5	101 (26)	101 (26)	57 (26)	25 (25)	20 (20)	57 (26)	H	26 (26)	43 (23)	18 (18)	20 (20)	0 / 92	9
Holy See	Europe	No data	101 (26)	101 (26)	0 (0)	0 (0)	0 (0)	0 (0)	No data	0 (0)	0 (0)	0 (0)	0 (0)	No data	No data
Italy	Europe	4	101 (26)	101 (26)	57 (26)	25 (25)	23 (23)	57 (26)	H	26 (26)	43 (23)	14 (14)	20 (20)	0 / 54	5
Kosovo	Europe	No data	0 (0)	0 (0)	57 (26)	25 (25)	0 (0)	35 (26)	LM	5 (5)	43 (23)	0 (0)	0 (0)	No data	No data
Malta	Europe	5	101 (26)	101 (26)	57 (26)	25 (25)	20 (20)	47 (26)	H	25 (25)	38 (18)	15 (15)	23 (23)	0 / 88	0
Montenegro	Europe	2	101 (26)	101 (26)	57 (26)	25 (25)	10 (10)	17 (16)	UM	10 (10)	0 (0)	2 (2)	0 (0)	No data	0
Portugal	Europe	5	101 (26)	101 (26)	57 (26)	25 (25)	24 (24)	57 (26)	H	26 (26)	43 (23)	17 (17)	19 (19)	0 / 64	11
San Marino	Europe	6	101 (26)	101 (26)	0 (0)	25 (25)	0 (0)	18 (17)	H	0 (0)	0 (0)	0 (0)	0 (0)	No data	0
Serbia	Europe	2	101 (26)	101 (26)	27 (26)	25 (25)	10 (10)	22 (21)	UM	25 (25)	0 (0)	10 (10)	7 (7)	No data	12
Slovenia	Europe	5	101 (26)	101 (26)	57 (26)	25 (25)	20 (20)	22 (21)	H	26 (26)	43 (23)	11 (11)	26 (26)	0 / 86	0
Spain	Europe	5	101 (26)	101 (26)	57 (26)	25 (25)	20 (20)	57 (26)	H	26 (26)	43 (23)	5 (5)	6 (6)	0 / 52	12
TFYR Mace-donia	Europe	2	101 (26)	101 (26)	57 (26)	25 (25)	9 (9)	27 (26)	UM	25 (25)	37 (17)	8 (8)	4 (4)	No data	No data
Austria	Europe	4	101 (26)	101 (26)	57 (26)	25 (25)	25 (25)	57 (26)	H	26 (26)	43 (23)	16 (16)	22 (22)	0 / 7	11
Belgium	Europe	4	101 (26)	101 (26)	57 (26)	25 (25)	21 (21)	57 (26)	H	26 (26)	43 (23)	19 (19)	20 (20)	0 / 12	5
France	Europe	4	101 (26)	101 (26)	57 (26)	25 (25)	22 (22)	57 (26)	H	26 (26)	43 (23)	16 (16)	20 (20)	No data	9
Germany	Europe	4	101 (26)	101 (26)	57 (26)	25 (25)	21 (21)	47 (26)	H	26 (26)	43 (23)	16 (16)	26 (26)	No data	22
Liechtenstein	Europe	7	101 (26)	101 (26)	0 (0)	25 (25)	22 (22)	46 (26)	H	0 (0)	0 (0)	22 (22)	19 (19)	No data	0

Table A.2 – Continued from previous page

Country	Region	Cl.	Tot. pop.	Urb. pop.	Y. pop.	El. acc.	Collected	GDP	GNI	En. cons.	Gh. gas	Collection	Recyc.	Lf, Dumps	Plastic
Luxembourg	Europe	6	101 (26)	101 (26)	57 (26)	25 (25)	25 (25)	57 (26)	H	26 (26)	43 (23)	19 (19)	19 (19)	0 / 19	1
Monaco	Europe	7	101 (26)	101 (26)	0 (0)	25 (25)	21 (21)	42 (22)	H	0 (0)	0 (0)	22 (22)	33 (33)	0 / 27	0
Netherlands	Europe	4	101 (26)	101 (26)	57 (26)	25 (25)	24 (24)	57 (26)	H	26 (26)	43 (23)	15 (15)	20 (20)	0 / 2	19
Switzerland	Europe	6	101 (26)	101 (26)	57 (26)	25 (25)	25 (25)	47 (26)	H	26 (26)	43 (23)	20 (20)	20 (20)	0 / 1	15
Anguilla	North America	No data	101 (26)	101 (26)	0 (0)	0 (0)	8 (8)	0 (0)	No data	0 (0)	0 (0)	8 (8)	9 (9)	No data	No data
Antigua and Barbuda	North America	3	101 (26)	101 (26)	57 (26)	25 (25)	11 (11)	40 (26)	H	5 (5)	31 (21)	11 (11)	15 (15)	0 / 99	0
Aruba	North America	3	101 (26)	101 (26)	57 (26)	25 (25)	0 (0)	18 (18)	H	0 (0)	18 (0)	0 (0)	0 (0)	No data	0
Bahamas, The	North America	3	101 (26)	101 (26)	57 (26)	25 (25)	3 (3)	57 (26)	H	5 (5)	33 (23)	0 (0)	0 (0)	No data	No data
Barbados	North America	3	101 (26)	101 (26)	57 (26)	15 (15)	0 (0)	27 (26)	H	5 (5)	34 (14)	0 (0)	0 (0)	No data	0
British Virgin Islands	North America	3	101 (26)	101 (26)	57 (26)	25 (25)	8 (8)	55 (25)	H	24 (24)	43 (23)	1 (1)	8 (8)	No data	0
Cayman Islands	North America	3	101 (26)	101 (26)	0 (0)	25 (25)	0 (0)	2 (2)	H	0 (0)	18 (2)	0 (0)	0 (0)	No data	0
Cuba	North America	3	101 (26)	101 (26)	57 (26)	25 (25)	14 (14)	46 (26)	UM	25 (25)	43 (23)	14 (14)	12 (12)	0 / 100	10
Dominica	North America	3	101 (26)	101 (26)	0 (0)	25 (25)	5 (5)	40 (26)	UM	5 (5)	43 (23)	5 (5)	0 (0)	0 / 100	0
Dominican Republic	North America	3	101 (26)	101 (26)	57 (26)	25 (25)	0 (0)	57 (26)	UM	25 (25)	43 (23)	9 (9)	0 (0)	No data	36
Grenada	North America	3	101 (26)	101 (26)	57 (26)	25 (25)	0 (0)	40 (26)	UM	5 (5)	43 (23)	0 (0)	0 (0)	0 / 90	0

Table A.2 – Continued from previous page

Country	Region	Cl.	Tot. pop.	Urb. pop.	Y. pop.	El. acc.	Collected	GDP	GNI	En. cons.	Gh. gas	Collection	Recyc.	Lf, Dumps	Plastic
Guadeloupe	North America	No data	101 (26)	101 (26)	0 (0)	0 (0)	6 (6)	0 (0)	No data	0 (0)	0 (0)	4 (4)	1 (1)	No data	No data
Haiti	North America	3	101 (26)	101 (26)	57 (26)	25 (25)	0 (0)	26 (25)	L	25 (25)	43 (23)	0 (0)	0 (0)	13 / 59	0
Jamaica	North America	3	101 (26)	101 (26)	57 (26)	25 (25)	7 (7)	57 (26)	UM	25 (25)	41 (23)	2 (2)	0 (0)	0 / 100	18
Martinique	North America	No data	101 (26)	101 (26)	0 (0)	0 (0)	6 (6)	0 (0)	No data	0 (0)	0 (0)	10 (10)	1 (1)	No data	No data
Montserrat	North America	No data	101 (26)	101 (26)	0 (0)	0 (0)	0 (0)	0 (0)	No data	0 (0)	0 (0)	0 (0)	0 (0)	No data	No data
Sint Maarten (Dutch part)	North America	No data	101 (26)	101 (26)	0 (0)	25 (25)	0 (0)	0 (0)	No data	0 (0)	0 (0)	0 (0)	0 (0)	No data	No data
Caribbean Netherlands	North America	No data	101 (26)	101 (26)	0 (0)	0 (0)	0 (0)	0 (0)	No data	0 (0)	0 (0)	0 (0)	0 (0)	No data	No data
Curaçao	North America	No data	101 (26)	101 (26)	57 (26)	25 (25)	0 (0)	0 (0)	H	25 (25)	0 (0)	0 (0)	0 (0)	No data	No data
Puerto Rico	North America	3	101 (26)	101 (26)	57 (26)	13 (13)	0 (0)	54 (24)	H	0 (0)	43 (23)	0 (0)	0 (0)	No data	0
Saint Kitts and Nevis	North America	3	101 (26)	101 (26)	0 (0)	25 (25)	0 (0)	57 (26)	H	5 (5)	20 (0)	0 (0)	0 (0)	No data	No data
Saint Lucia	North America	3	101 (26)	101 (26)	57 (26)	25 (25)	12 (12)	40 (26)	UM	5 (5)	41 (21)	12 (12)	8 (8)	No data	No data
St. Martin (French part)	North America	No data	0 (0)	0 (0)	0 (0)	25 (25)	0 (0)	0 (0)	No data	0 (0)	0 (0)	0 (0)	0 (0)	No data	No data

Table A.2 – Continued from previous page

Country	Region	Cl.	Tot. pop.	Urb. pop.	Y. pop.	El. acc.	Collected	GDP	GNI	En. cons.	Gh. gas	Collection	Recyc.	Lf, Dumps	Plastic
Saint Vincent and the Grenadines	North America	3	101 (26)	101 (26)	57 (26)	25 (25)	1 (1)	27 (26)	UM	24 (24)	43 (23)	1 (1)	2 (2)	No data	No data
Trinidad and Tobago	North America	3	101 (26)	101 (26)	57 (26)	25 (25)	9 (9)	57 (26)	H	25 (25)	16 (14)	1 (1)	0 (0)	43 / 0	24
Turks and Caicos Islands	North America	No data	101 (26)	101 (26)	0 (0)	25 (25)	0 (0)	0 (0)	H	0 (0)	25 (5)	0 (0)	0 (0)	No data	0
United States Virgin Islands	North America	3	101 (26)	101 (26)	0 (0)	0 (0)	0 (0)	0 (0)	H	0 (0)	22 (3)	0 (0)	0 (0)	No data	No data
Belize	North America	2	101 (26)	101 (26)	57 (26)	25 (25)	2 (2)	57 (26)	UM	5 (5)	43 (23)	16 (16)	0 (0)	0 / 100	5
Costa Rica	North America	3	101 (26)	101 (26)	57 (26)	25 (25)	1 (1)	57 (26)	UM	25 (25)	43 (23)	1 (1)	0 (0)	22 / 72	18
El Salvador	North America	2	101 (26)	101 (26)	57 (26)	25 (25)	0 (0)	52 (26)	LM	25 (25)	43 (23)	0 (0)	0 (0)	No data	0
Guatemala	North America	2	101 (26)	101 (26)	57 (26)	25 (25)	0 (0)	57 (26)	LM	25 (25)	43 (23)	4 (4)	0 (0)	49 / 11	13
Honduras	North America	1	101 (26)	101 (26)	57 (26)	25 (25)	3 (3)	57 (26)	LM	25 (25)	43 (23)	3 (3)	0 (0)	No data	0
Mexico	North America	3	101 (26)	101 (26)	57 (26)	25 (25)	22 (22)	57 (26)	UM	26 (26)	43 (23)	2 (2)	18 (18)	0 / 97	6
Nicaragua	North America	1	101 (26)	101 (26)	57 (26)	25 (25)	0 (0)	28 (26)	LM	25 (25)	43 (23)	0 (0)	0 (0)	34 / 28	0
Panama	North America	3	101 (26)	101 (26)	57 (26)	25 (25)	21 (21)	57 (26)	UM	25 (25)	43 (23)	8 (8)	0 (0)	20 / 56	11

Table A.2 – Continued from previous page

Country	Region	Cl.	Tot. pop.	Urb. pop.	Y. pop.	El. acc.	Collected	GDP	GNI	En. cons.	Gh. gas	Collection	Recyc.	Lf, Dumps	Plastic
Argentina	South America	3	101 (26)	101 (26)	57 (26)	25 (25)	5 (5)	55 (26)	UM	25 (25)	43 (23)	2 (2)	0 (0)	No data	14
Bolivia (Plurinational State of)	South America	1	101 (26)	101 (26)	57 (26)	25 (25)	17 (17)	57 (26)	LM	25 (25)	43 (23)	11 (11)	0 (0)	No data	No data
Brazil	South America	3	101 (26)	101 (26)	57 (26)	25 (25)	11 (11)	57 (26)	UM	25 (25)	43 (23)	18 (18)	1 (1)	No data	15
Chile	South America	3	101 (26)	101 (26)	57 (26)	25 (25)	12 (12)	57 (26)	H	26 (26)	43 (23)	0 (0)	0 (0)	0 / 100	10
Colombia	South America	2	101 (26)	101 (26)	57 (26)	25 (25)	9 (9)	57 (26)	UM	25 (25)	43 (23)	6 (6)	7 (7)	54 / 46	10
Ecuador	South America	2	101 (26)	101 (26)	57 (26)	25 (25)	2 (2)	57 (26)	UM	25 (25)	43 (23)	9 (9)	2 (2)	No data	0
Falkland Islands (Malvinas)	South America	No data	101 (26)	101 (26)	0 (0)	0 (0)	0 (0)	0 (0)	No data	0 (0)	0 (0)	0 (0)	0 (0)	No data	No data
French Guiana	South America	No data	101 (26)	101 (26)	0 (0)	0 (0)	9 (9)	0 (0)	No data	0 (0)	0 (0)	3 (3)	0 (0)	No data	No data
Guyana	South America	2	101 (26)	101 (26)	57 (26)	25 (25)	0 (0)	57 (26)	UM	5 (5)	43 (23)	0 (0)	0 (0)	37 / 59	10
Paraguay	South America	2	101 (26)	101 (26)	57 (26)	25 (25)	0 (0)	52 (26)	UM	25 (25)	43 (23)	18 (18)	0 (0)	42 / 44	0
Peru	South America	2	101 (26)	101 (26)	57 (26)	25 (25)	1 (1)	57 (26)	UM	25 (25)	43 (23)	6 (6)	1 (1)	19 / 66	4
Suriname	South America	3	101 (26)	101 (26)	57 (26)	25 (25)	9 (9)	57 (26)	UM	15 (15)	41 (21)	3 (3)	0 (0)	100 / 0	0

Table A.2 – Continued from previous page

Country	Region	Cl.	Tot. pop.	Urb. pop.	Y. pop.	El. acc.	Collected	GDP	GNI	En. cons.	Gh. gas	Collection	Recyc.	Lf, Dumps	Plastic
Uruguay	South America	3	101 (26)	101 (26)	57 (26)	25 (25)	3 (3)	30 (26)	H	25 (25)	43 (23)	0 (0)	3 (3)	32 / 3	11
Venezuela (Bolivarian Republic of)	South America	3	101 (26)	101 (26)	57 (26)	25 (25)	0 (0)	57 (26)	UM	5 (5)	42 (22)	0 (0)	0 (0)	No data	No data
Bermuda	North America	6	101 (26)	101 (26)	0 (0)	25 (25)	12 (12)	54 (24)	H	0 (0)	37 (20)	12 (12)	12 (12)	No data	0
Canada	North America	4	101 (26)	101 (26)	57 (26)	25 (25)	0 (0)	57 (26)	H	26 (26)	43 (23)	1 (1)	0 (0)	0 / 42	3
Greenland	North America	4	101 (26)	101 (26)	0 (0)	25 (25)	0 (0)	46 (26)	H	0 (0)	0 (0)	0 (0)	0 (0)	No data	0
Saint Pierre and Miquelon	North America	No data	101 (26)	101 (26)	0 (0)	0 (0)	0 (0)	0 (0)	No data	0 (0)	0 (0)	0 (0)	0 (0)	No data	No data
United States of America	North America	4	101 (26)	101 (26)	57 (26)	25 (25)	23 (23)	57 (26)	H	25 (25)	43 (23)	2 (2)	19 (19)	No data	No data
Australia	Oceania	4	101 (26)	101 (26)	57 (26)	25 (25)	4 (4)	57 (26)	H	26 (26)	43 (23)	0 (0)	3 (3)	0 / 70	4
New Zealand	Oceania	4	101 (26)	101 (26)	57 (26)	25 (25)	11 (11)	56 (26)	H	26 (26)	43 (23)	0 (0)	0 (0)	0 / 85	8
Fiji	Oceania	2	101 (26)	101 (26)	57 (26)	25 (25)	0 (0)	57 (26)	UM	5 (5)	43 (23)	0 (0)	0 (0)	No data	8
New Caledonia	Oceania	2	101 (26)	101 (26)	57 (26)	20 (20)	0 (0)	36 (11)	H	0 (0)	9 (4)	0 (0)	0 (0)	No data	0
Papua New Guinea	Oceania	2	101 (26)	101 (26)	57 (26)	25 (25)	0 (0)	57 (26)	LM	0 (0)	43 (23)	0 (0)	0 (0)	No data	0
Solomon Islands	Oceania	2	101 (26)	101 (26)	57 (26)	23 (23)	0 (0)	49 (26)	LM	5 (5)	43 (23)	0 (0)	0 (0)	No data	17
Vanuatu	Oceania	2	101 (26)	101 (26)	57 (26)	25 (25)	0 (0)	32 (26)	LM	24 (24)	43 (23)	0 (0)	0 (0)	No data	8
Guam	Oceania	2	101 (26)	101 (26)	57 (26)	20 (20)	0 (0)	15 (14)	H	0 (0)	43 (23)	0 (0)	0 (0)	No data	0

Table A.2 – Continued from previous page

Country	Region	Cl.	Tot. pop.	Urb. pop.	Y. pop.	El. acc.	Collec-ted	GDP	GNI	En. cons.	Gh. gas	Collec-tion	Recyc.	Lf, Dumps	Plastic
Kiribati	Oceania	2	101 (26)	101 (26)	57 (26)	25 (25)	0 (0)	47 (26)	LM	5 (5)	27 (17)	0 (0)	0 (0)	No data	0
Marshall Islands	Oceania	2	101 (26)	101 (26)	0 (0)	25 (25)	1 (1)	36 (26)	UM	4 (4)	20 (0)	1 (1)	1 (1)	33 / 59	15
Micronesia (Fed. States of)	Oceania	2	101 (26)	101 (26)	57 (26)	25 (25)	0 (0)	32 (26)	No data	0 (0)	27 (7)	0 (0)	0 (0)	No data	No data
Nauru	Oceania	2	101 (26)	101 (26)	0 (0)	15 (15)	0 (0)	10 (9)	H	0 (0)	0 (0)	0 (0)	0 (0)	No data	0
Northern Mariana Islands	Oceania	2	101 (26)	101 (26)	0 (0)	20 (20)	0 (0)	15 (14)	H	0 (0)	43 (23)	0 (0)	0 (0)	No data	0
Palau	Oceania	2	101 (26)	101 (26)	0 (0)	25 (25)	0 (0)	27 (26)	UM	1 (1)	0 (0)	0 (0)	0 (0)	No data	0
American Samoa	Oceania	2	101 (26)	101 (26)	0 (0)	0 (0)	0 (0)	15 (14)	UM	0 (0)	43 (23)	0 (0)	0 (0)	No data	0
Cook Islands	Oceania	No data	101 (26)	101 (26)	0 (0)	0 (0)	0 (0)	0 (0)	No data	0 (0)	0 (0)	0 (0)	0 (0)	No data	No data
French Polynesia	Oceania	2	101 (26)	101 (26)	57 (26)	25 (25)	1 (1)	36 (11)	H	0 (0)	21 (11)	0 (0)	0 (0)	No data	0
Niue	Oceania	No data	101 (26)	101 (26)	0 (0)	0 (0)	0 (0)	0 (0)	No data	0 (0)	0 (0)	0 (0)	0 (0)	No data	No data
Samoa	Oceania	2	101 (26)	101 (26)	57 (26)	25 (25)	0 (0)	38 (26)	LM	5 (5)	43 (23)	0 (0)	0 (0)	No data	0
Tokelau	Oceania	No data	101 (26)	101 (26)	0 (0)	0 (0)	0 (0)	0 (0)	No data	0 (0)	0 (0)	0 (0)	0 (0)	No data	No data
Tonga	Oceania	2	101 (26)	101 (26)	57 (26)	25 (25)	0 (0)	42 (26)	LM	5 (5)	42 (22)	0 (0)	0 (0)	No data	5
Tuvalu	Oceania	No data	101 (26)	101 (26)	0 (0)	0 (0)	0 (0)	0 (0)	UM	0 (0)	0 (0)	0 (0)	0 (0)	No data	No data
Wallis and Futuna Islands	Oceania	No data	101 (26)	101 (26)	0 (0)	0 (0)	0 (0)	0 (0)	No data	0 (0)	0 (0)	0 (0)	0 (0)	No data	No data

7.11 Predictor Variable Distribution of Erroneous Countries

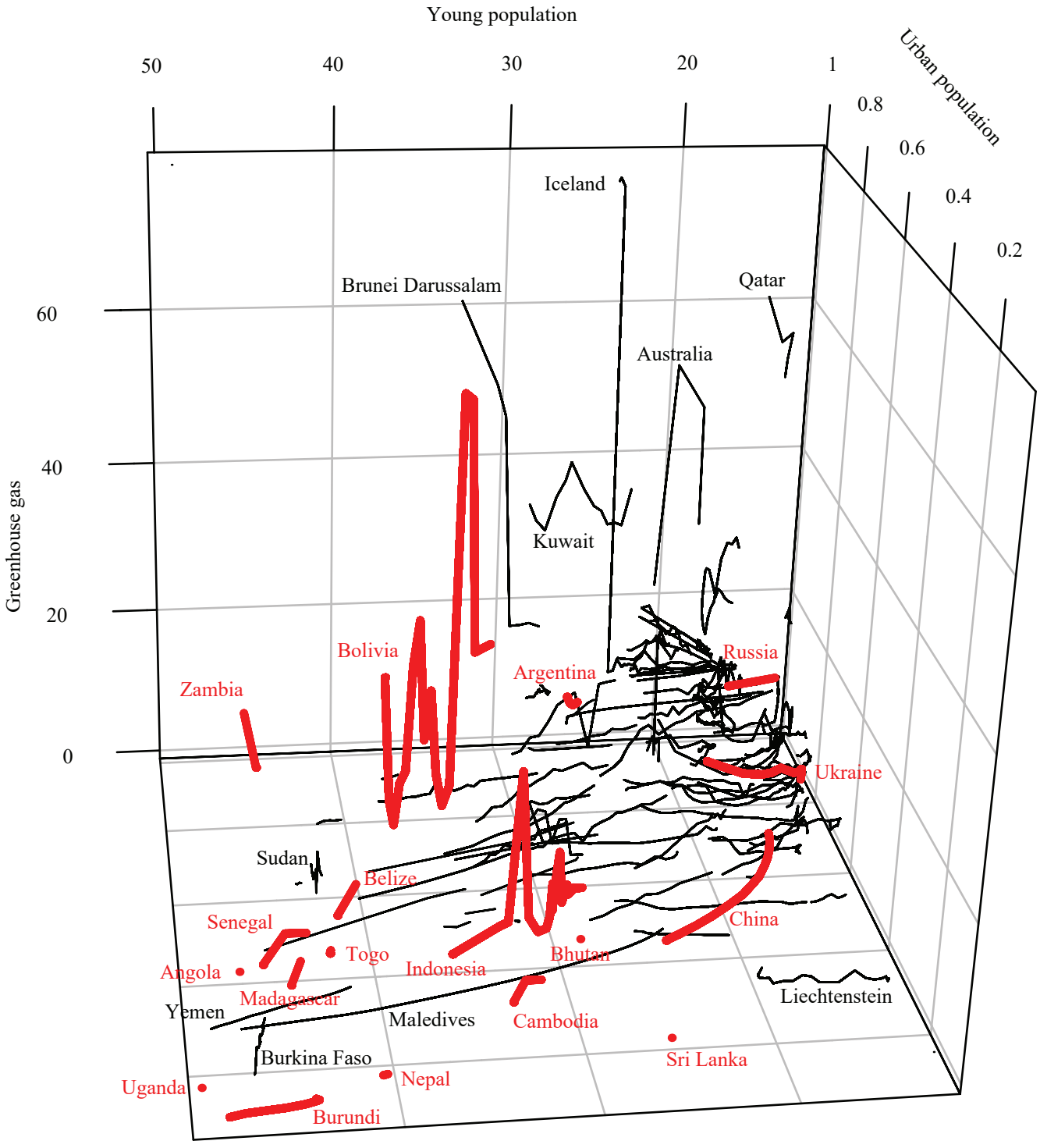

Figure A.7. Three dimensional representation of the three predictors that were used in the NNET and RF models to estimate $MSW_{collected}$. The countries in red have an prediction error >30%. The prediction of countries with a high fraction of population <14 years and a low fraction of urban population or with a steep gradient between these variables is often erroneous. Only three countries do not show these traits: Argentina Russia and Ukraine

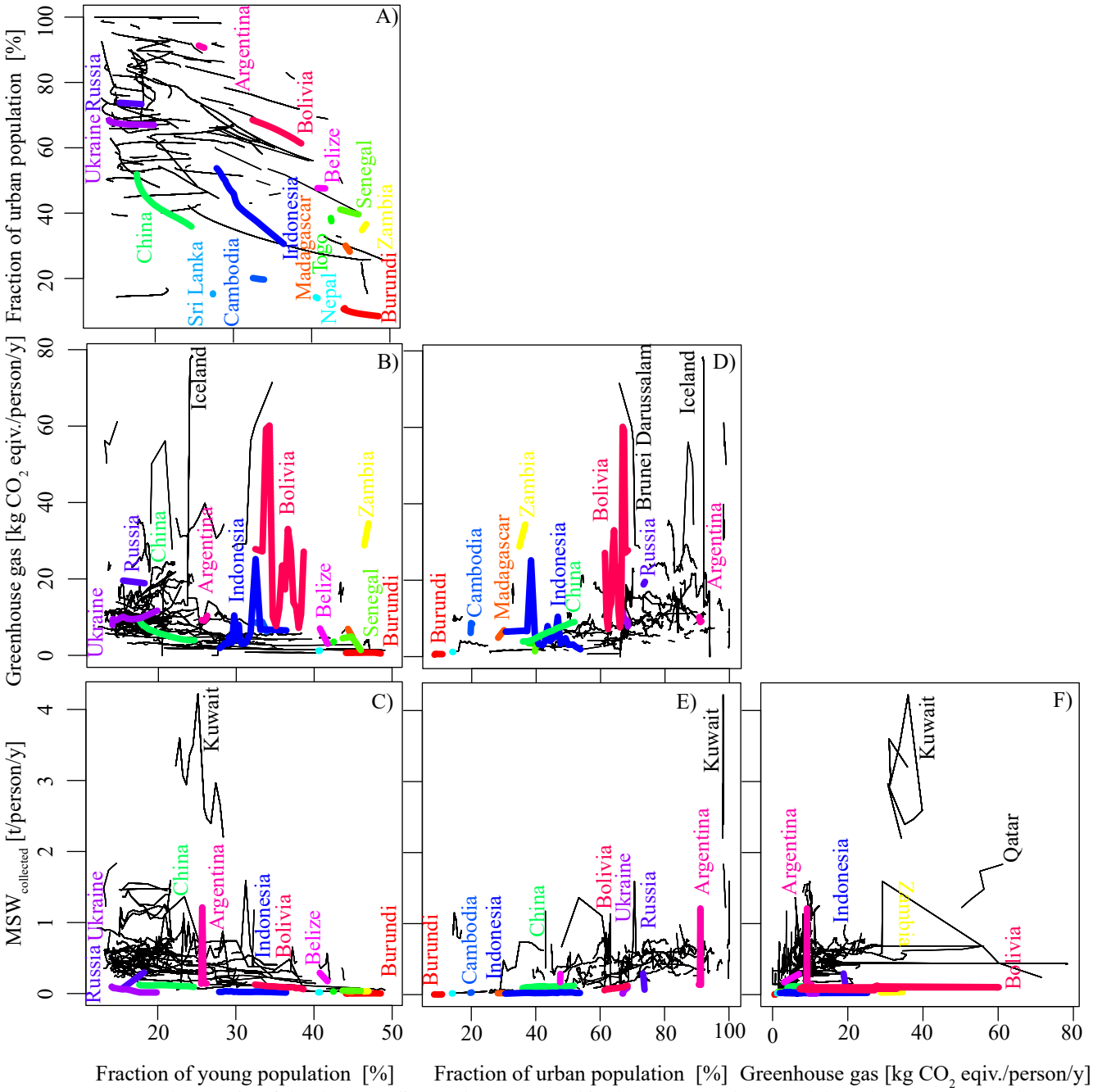

Figure A.8. Scatterplot matrix of the three predictors that were used in the NNET and RF models and $MSW_{collected}$. The coloured lines represent countries countries that have a prediction error >30%.

7.12 Predictions of P_O from coastal sources and in total

Table A.3. Summary of the modelled coastal plastic input to the ocean in 1990 and 2015 for the income classes, the continents and worldwide. See section 3.3.2 for more details.

	Subset	Total marine plastic input [kt/y]	
		1990	2015
Income	High-income	1'299	1'640
	Upper middle-income	9'318	9'762
	Lower middle-income	2'309	6'117
	Low-income	44	173
Continents	Africa	279	728
	Asia	11'2162	15'015
	Europe	900	733
	North America & Caribbean	331	587
	South America	213	593
	Oceania	29	36
	World	12'972	17'693

Table A.4. Summary of the modelled total plastic input to the ocean in 1990 and 2015 for the income classes, the continents and worldwide. See section 3.3.2 for more details.

	Subset	Total marine plastic input [kt/y]	
		1990	2015
Income	High-income	1'372	1'735
	Upper middle-income	9'410	10'054
	Lower middle-income	2'604	6'755
	Low-income	66	208
Continents	Africa	340	864
	Asia	11'442	15'754
	Europe	1'063	847
	North America & Caribbean	359	637
	South America	220	615
	Oceania	31	37
	World	13'455	18'754

Eidgenössische Technische Hochschule Zürich
Swiss Federal Institute of Technology Zurich

Declaration of originality

The signed declaration of originality is a component of every semester paper, Bachelor's thesis, Master's thesis and any other degree paper undertaken during the course of studies, including the respective electronic versions.

Lecturers may also require a declaration of originality for other written papers compiled for their courses.

I hereby confirm that I am the sole author of the written work here enclosed and that I have compiled it in my own words. Parts excepted are corrections of form and content by the supervisor.

Title of work (in block letters):

Spatio-Temporal Estimation of Global Plastic Waste Inputs to the Ocean from 1990 to 2015

Authored by (in block letters):

For papers written by groups the names of all authors are required.

Name(s):	**First name(s):**
Lang	Kevin

With my signature I confirm that

- I have committed none of the forms of plagiarism described in the 'Citation etiquette' information sheet.
- I have documented all methods, data and processes truthfully.
- I have not manipulated any data.
- I have mentioned all persons who were significant facilitators of the work.

I am aware that the work may be screened electronically for plagiarism.

Place, date	**Signature(s)**
Aarau, 24.09.2018	

For papers written by groups the names of all authors are required. Their signatures collectively guarantee the entire content of the written paper.