Early estimates of environmental accounts (economy-wide material flow accounts)

A. de la Fuente[1], R. Marra Campanale[2]

[1] *Eurostat, Luxembourg; Arturo.De-La-Fuente@ec.europa.eu*
[2] *Eurostat, Luxembourg; Renato.Marra-Campanale@ec.europa.eu*

Abstract
Environmental accounting is a multipurpose conceptual framework that describes the interactions between the economy and the environment. After key EU policy users called for an improvement in timeliness, Eurostat set up production of early estimates in one module of the European environmental accounts, economy-wide material flow accounts, for the EU and Member States at T+9 months and T+18 months. Eurostat bases the T+18 estimates on national and European data sources and the T+9 estimates on regression modelling techniques and predictor variables. The use of modelling in official statistics raises some questions. Some argue that official statistics should not make use of models. Actually, modelling is everywhere as almost any step of statistical production requires some form of model or assumption. Requirements are the use of well-established and well-tested models, proper documentation of methods, quality assessment and transparency towards users.

Keywords: environmental accounts, timeliness, data modelling.

1. Context

Timeliness is arguably the most straightforward quality criterion[1] for users to grasp: a look at a statistical database tells how old are the most recent data available. This shapes at a glance the users' perception of the data's statistical quality.

It is well known that timeliness can be improved at the expense of other quality criteria, typically accuracy. The particularities depend on the context. Timeliness is constrained because most statistical operations require that the reporting units have at hand certain

[1] The quality criteria for European statistics are defined in Regulation 223/2009, Article 12

information they need for the statistical reporting, e.g. businesses to have their quarterly/annual bookkeeping to report in business statistics; households to have completed their income tax returns to report in income statistics, etc. Generally speaking, the availability of this underlying information is the most important single factor to determine how timely statistics can be. Other factors exist too, e.g. statistical operations take some time to compile the information, ensure validation, publication and transmission to international organisations, etc. but generally those are not the main reason for the timeliness of the statistics.

Whenever statistics are required to improve their timeliness and the necessary information is not available yet, it is possible to take recourse to alternative techniques, e.g. based on modelling. This is very much used in the economic domain, where flash estimates, nowcasts and forward estimates are standard practice. Such practices are less established in other statistical domains, such as social and environmental statistics, but pressure is growing to provide more up to date data for monitoring policies.

This paper reports the Eurostat experience to improve the timeliness of environmental accounts, in particular the so-called economy-wide material flow accounts. This has been done with the release of early estimates at 18 months and 9 months after the reference period (hereinafter, at 'T+18 months' and 'T+9 months') in addition to the normal release at T+27 months. As a result the timeliness of those accounts has improved by 18 months.

2. Environmental accounts

Environmental-economic accounts (environmental accounts for short) are a statistical system bringing together economic and environmental information in a common framework to measure the contribution of the environment to the economy and the impact of the economy on the environment. Environmental accounts offer a means of monitoring the pressures exerted by the economy on the environment and of exploring how these might be abated. They organise environmental data from many domains using the same concepts and terminology as the national accounts, and thus they show the interaction between economic, household and environmental factors and consequently are more informative than national accounts alone.

In Europe, environmental accounts are produced within the framework of the international standard System of Environmental-Economic Accounting 2012 – Central Framework and established in the Regulation (EU) No. 691/2011. The areas covered in the Regulation are:

- Emissions to air (greenhouse gases and air pollutants);
- Material flows at economy-wide level;
- Physical energy flows;
- Environmental taxes;
- Environmental protection expenditure;
- Environmental goods and services sector (gross value added, employment, etc.).

These six areas are developed in separate and independent modules. This paper focuses on the economy-wide material flow accounts (EW-MFA). Those accounts trace the amounts of materials entering the economy. Domestic extractions and imports of materials are fed to the economy, which processes them as outputs for domestic use or export, as well as generating residuals (emissions, etc.). EW-MFA can be used to derive indicators on resource extraction, material consumption and footprint of economies as well as their resource productivity in order to display whether decoupling between economic activity and use of natural resources occurs (i.e., growth can be maintained with fewer resources).

EW-MFA is probably the environmental accounts module with the highest EU policy profile, because it is used for the headline indicator resource productivity of the flagship initiative A resource-efficient Europe under the Europe 2020 Strategy. This initiative supports the shift towards a resource-efficient, low-carbon economy to achieve sustainable growth.

3. EW-MFA early estimates

EW-MFA estimates are annual, the same as the other environmental accounts. According to the Regulation (EU) No. 691/2011, the statistical authorities in the Member States must produce the EW-MFA estimates and transmit them to Eurostat within 24 months of the end of the reference year, i.e. at T+24 months. This legal provision determines the timeliness of the

data published by Eurostat. After data validation and calculation of EU aggregates, Eurostat publishes the results at T+27 months.

For a number of years key EU policy users called for an improvement of this timeliness for economic analysis, decision-taking and policy-making. The user demand induced Eurostat to create in 2014 early estimates at T+18 months and T+9 months on top of the standard EW-MFA production cycle (see Table 1). These are early estimates for each of the 28 EU Member States and EU totals.

Table 1: Overview of Eurostat EW-MFA releases scheduled for 2016

	Standard release	**Early estimates**	
Release	T+27	T+18	T+9
Date of release	March 2016	June 2016	September 2016
Reference year	2013	2014	2015
Summary method	Legal country transmission in December 2015 (i.e., T+24)	Based on voluntary data transmitted in December 2015 plus available data sources	Based on T+27 and T+18 data at Eurostat plus modelling

The main strength of the early estimates at T+18 is that they enable Eurostat to release one extra year of data only a few months after the standard release by lifting the constraint to publish only data transmitted by the countries. Early estimates at T+9 push this principle further by adding one year more, with fewer breakdowns and higher errors.

The early estimates at T+18 and T+9 have the same scope (countries, variables) but different material category breakdowns: estimates at T+18 have the full breakdowns (more than 50 materials) whereas estimates at T+9 are only produced for 4 main categories (biomass, metal ores, non-metallic minerals and fossil energy materials). They also differ in the method applied for producing data for the indicator domestic extraction. The next sections briefly explain the estimation methods.

3.1. Estimation method for early estimates at T+18

As stated above, Regulation (EU) No. 691/2011 requires countries to transmit data at T+24 months. Eurostat further invites countries to voluntarily transmit in the same questionnaire one additional year of data; however, not all countries are able to do it due to availability of data sources. For instance, in the data transmission with reporting deadline December 2013 the legal obligation was to transmit data up to 2011; more than half of the countries (16 out of 27) were able to deliver one extra year. This was the basis for the early estimate at T+18 months of the remaining 11 countries.[2]

This estimation involves a good deal of manual work and expert analysis. Domestic extraction is estimated separately from physical imports and exports. Estimating domestic extraction requires finding and exploiting national data sources, many of which are available online (e.g. yearbooks, national statistical institutes, ministries, geological surveys and energy balances). Whenever this is not possible, other EU or international data sources are used as second best source, e.g. the Eurostat online database, US and British Geological Survey. Most of these data are provisional, hence subject to revision, at the time when Eurostat undertakes this work. The estimation of physical trade is based on Eurostat's foreign trade statistics. Those figures are often subject to revision too.

Afterwards Eurostat communicates the estimates to the countries concerned and seeks their agreement before publication. Figures are published flagged as Eurostat estimates. Once estimates for all the countries are available, EU aggregates are produced bottom up.

3.2. Estimation method for early estimates at T+9

Eurostat produces early estimates at T+9 for the mandatory tables of the questionnaire – domestic extraction and trade tables – at the level of the four main material categories. Domestic extraction is estimated separately from physical trade. For physical imports and exports the same source and method are used as in the T+18 estimates, i.e. Eurostat's foreign

[2] Belgium, Cyprus, Estonia, Greece, Hungary, Ireland, Luxembourg, Malta, Netherlands, Spain, UK. These countries accounted for some 25% of the EU-27 prominent policy indicator *domestic material consumption*.

trade statistics. For domestic extraction the method is based on different predictive models for the four main material categories. The estimates for each combination of country and material are based on a model which is selected among those illustrated in Table 2.

Table 2: Predictive models available for domestic extraction (DE)[3]

Regression of absolute values	$\boldsymbol{DE^p_{(t)} = a + bP_{(t)}}$
Regression of absolute changes	$\boldsymbol{DE^p_{(t)} = DE_{(t-1)} + \{a + b[P_{(t)} - P_{(t-1)}]\}}$
Regression of relative changes	$\boldsymbol{DE^p_{(t)} = DE_{(t-1)} + 1 - \{a + b[P_{(t)} - P_{(t-1)}]/P_{(t-1)}\}}$
Annual relative change in predictor	$\boldsymbol{DE^p_{(t)} = DE_{(t-1)}\, P_{(t)}/P_{(t-1)}}$
Sum of individual fuels (fossil energy materials only)	$\boldsymbol{DE^p_{(t)} = \sum i \left[DE_{(t-1)}\, F_{i_{(t)}}/F_{i_{(t-1)}}\right]}$
Value of previous year	$\boldsymbol{DE^p_{(t)} = DE_{(t-1)}}$

More than thirty predictors feed these models.[4] Models and predictors are integrated in a tailor-made IT tool that allows comparing the chosen specification with the outcome from other potential specifications.

After estimating all countries, EU aggregates are produced bottom up.

[3] Where: *P* is the Predictor; DE^p stands for DE predicted; F_i is the fuel. The first three models are all estimated using OLS regressions.

[4] Production value of *Crop*; Gross value added in *Agriculture, forestry and fishing*; Volume index of production of *Manufacture of food products; beverages and tobacco products*; Production value of *Agriculture goods*; Volume index of production of *Processing and preserving of fish, crustaceans and molluscs*; Production value of *Forage plants*; Production value of *Live bovine animals*; Volume index of production of *Manufacture of wood and of products of wood and cork, except furniture, manufacture of articles of straw and plaiting materials*; Volume index of production of *Mining of iron ores*; Volume index of production of *Mining of metal ores*; Volume index of production of *Other mining and quarrying*; Gross value added in *Construction*; Volume index of production of *Construction*; Volume index of production of *Mining of coal and lignite*; Extraction of crude petroleum and natural gas; Primary production of natural gas; Primary production of hard coal; Primary production of lignite/brown coal; Primary production of peat; Primary production of crude oil (without NGL); Primary production of natural gas liquids (NGL); Volume index of production of *Extraction of crude petroleum.*

4. Quality assessment

This section assesses the quality of the early estimates at T+9 in terms of size of revisions, existence of systemic errors due to data revisions and errors due to the model used. The early estimates at T+18 are not analysed due to lack of space.

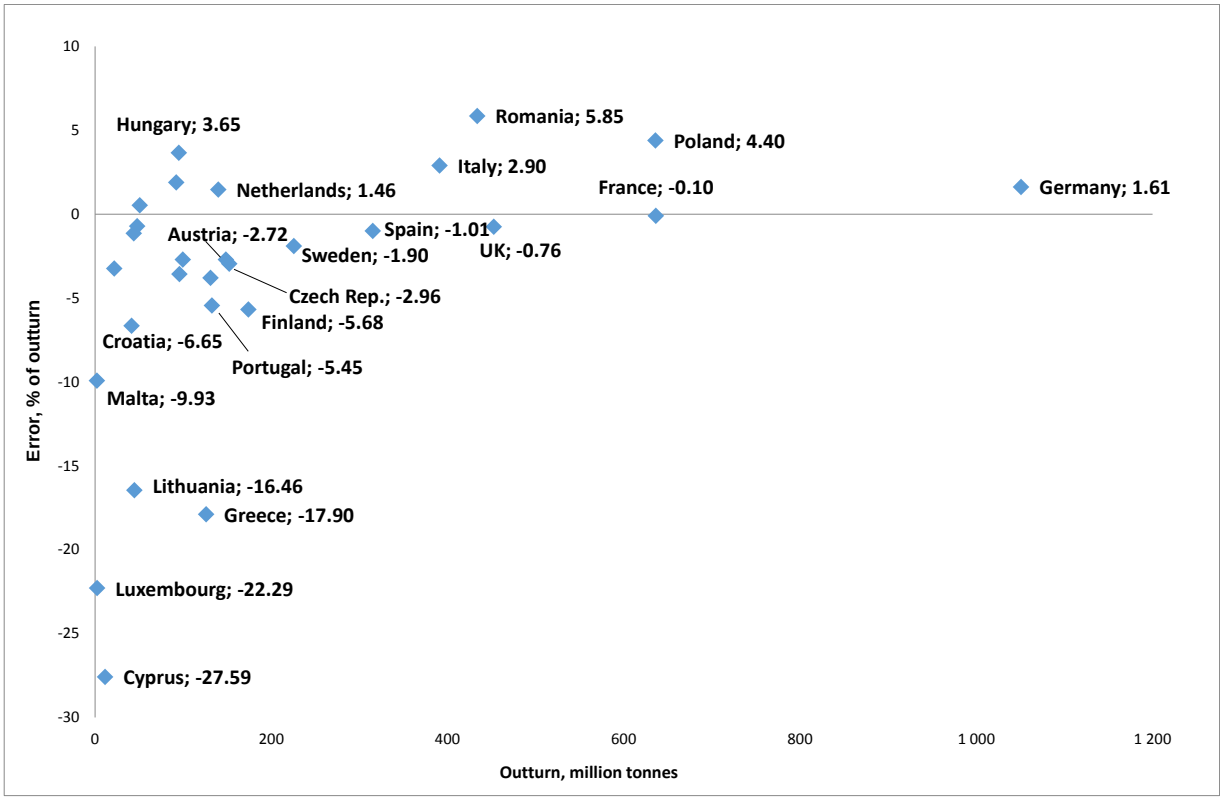
The size of the revisions is assessed comparing the outturns of domestic extraction for the reference year 2013 estimated at T+9 months (in September 2014) with the initial outturns estimated at T+18 months (in June 2015). Table 3 provides an overview of the size of the revisions for the EU-28 early estimates by material category. Only a 0.05 % revision is accounted for the EU-28 level and for total materials. The revisions by material category are bigger but they must be considered in the light of their size out of the total, e.g. metal ores.

Table 3: EU-28 early estimates of domestic extraction at T+9 for 2013 and revision size *(million tonnes and percentages)*

	Total	Biomass	Metal ores	Non-metallic minerals	Fossil fuels
Outturns (million tonnes)	5 795.5	1 710.8	196.3	3 129.8	758.5
Revision size (%)	0.05	-0.4	-11.9	1.3	-0.9

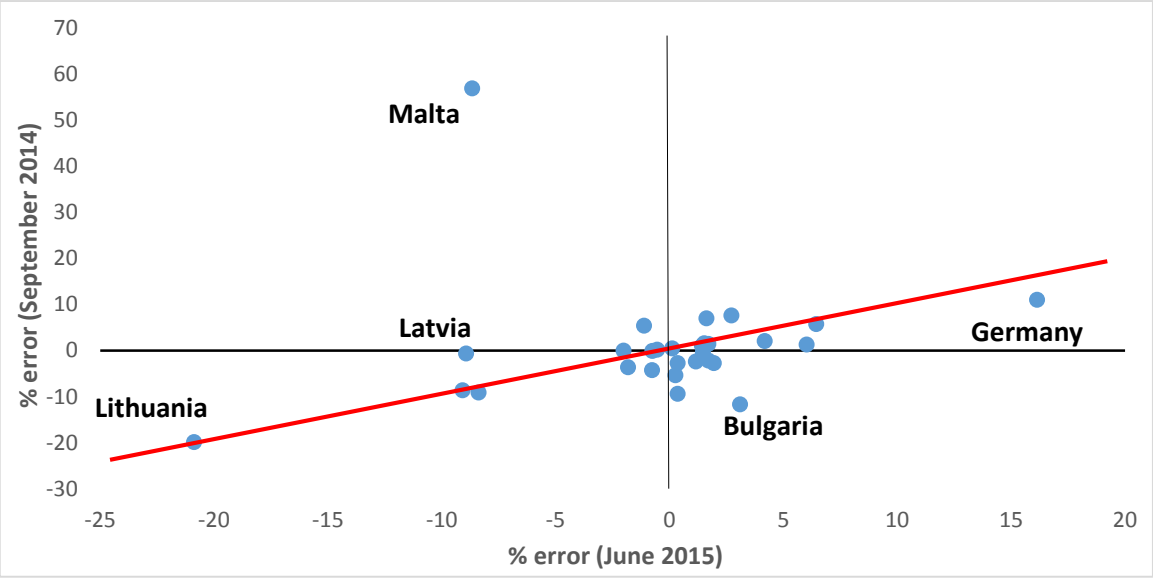
The revision of the EU estimate, indeed very small, stems from a composite range of more detailed estimates by country as shown in Figure 1. This figure shows that countries that account for a higher share of domestic extraction, e.g. Germany, France, Poland, UK, Romania and Italy have much more reliable early estimates.

Figure 1: Early estimates of domestic extraction by country at T+9 for 2013 and revision size *(thousand tonnes and percentages)*



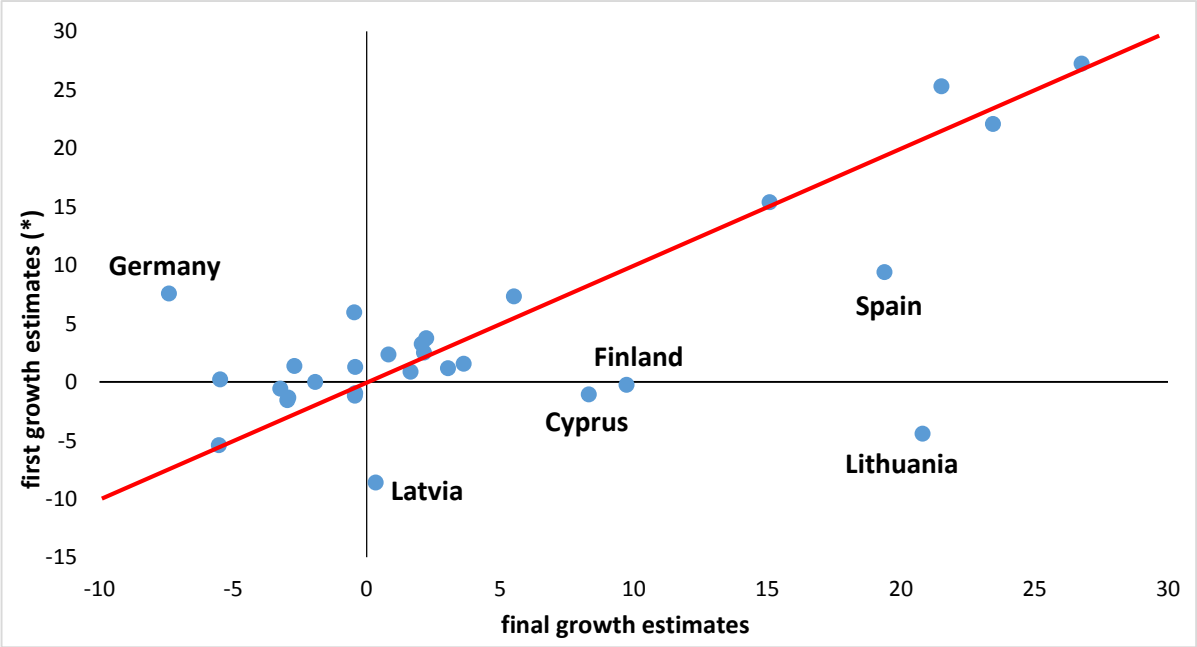
A second quality aspect analysed is the possible existence of a systemic error, i.e., a clear pattern of overestimation or underestimation, due to the revision of the underlying data between September 2014 and June 2015. Figure 2 illustrates it for domestic extractions of the material category biomass. The X-axis shows the size of the revision stemming from the models using the data available in September 2014. The Y-axis shows the size of the revision stemming from the models using the data available in June 2015. For countries over the red line, the new data does not affect the size of the errors. Since the observations are randomly distributed around (0,0) without any pattern, we conclude that the new information between September 2014 and June 2015 had no systematic impact on the estimation errors.

Figure 2: Comparison of errors in early estimates of domestic extraction for biomass due to different underlying data vintages. Year 2013. *(percentages)*



A third aspect investigated is the error due to the model selected. This analysis compares two estimates of domestic extraction growth between 2012-13: one is the growth between the first 2013 early estimate and the final 2012 estimate (see Figure 3, Y-axis); the other one is the growth between the final estimates for 2013 and 2012 (see Figure 3, X-axis). In order to neutralise the influence in first 2013 early estimate of revisions of the underlying data (already analysed above), the early estimate was re-calculated incorporating the latest data. Figure 3 presents this comparison for domestic extraction of biomass. The red line indicates the points where the first selection of the T+9 model got confirmed in the T+18 estimation. Figure 3 shows for biomass that the methods would have mostly predicted the correct direction of change 2012-2013 because the points are generally distributed around the red line. Among the instances where the early estimate mispredicts the direction of growth, e.g. Germany and Lithuania, only a different specification of the model would have better predicted the outturn as this part of the assessment excludes errors due to data revisions that are unrelated to the process of producing the early estimates.

Figure 3: 2012-13 growth of domestic extraction for biomass: first estimate(*) vs final estimate *(percentages)*



(*) re-calculated incorporating the final data

5. Costs and benefits

The quality assessment in the previous section proves to be satisfactory. Two further aspects to be considered are the costs and benefits of the early estimates.

There are some costs associated to the production of early estimates on the top of the normal EW-MFA production cycle. Fortunately hardly any additional data validation is necessary for the early estimates because there are no additional country transmissions besides those established in the Regulation. The additional tasks are rather as follows: for the estimate at T+18, manually producing country estimates for some countries; for the estimate at T+9, feeding the IT model tool with predictor data and selecting the best models. Part of this work is done by Eurostat staff and part is contracted out from Eurostat budget. In addition, for both releases there is modest workload related to database management and dissemination, performed by Eurostat. All in all, the annual work is estimated at 12 person-days of Eurostat staff plus 30 person-days of the contractor company.

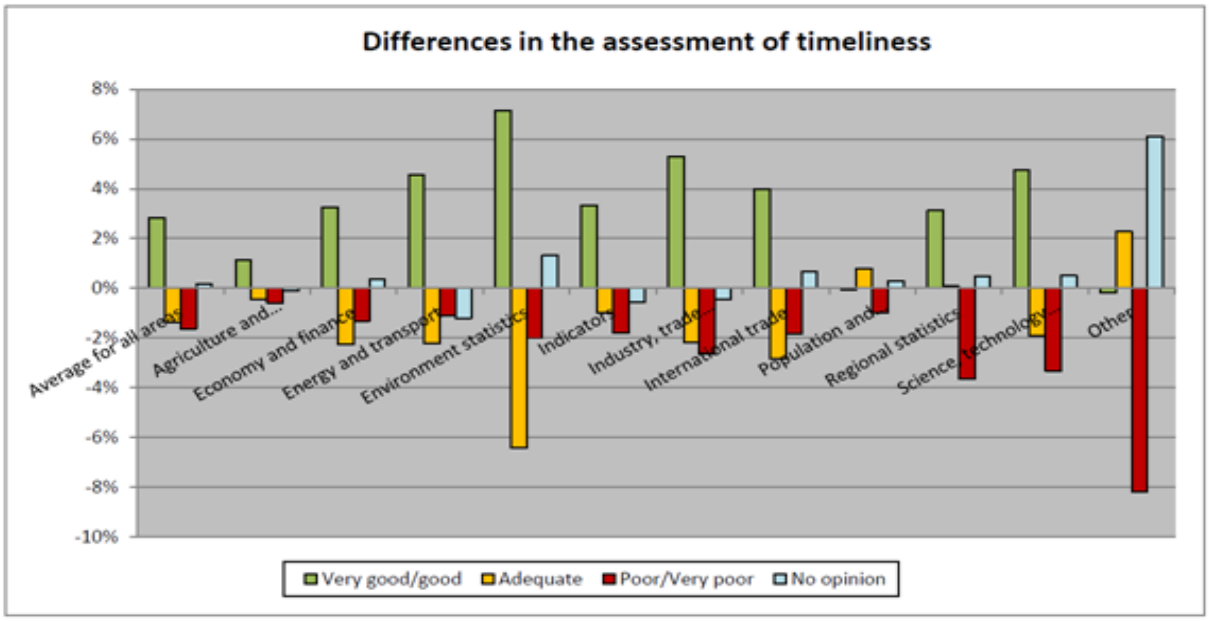
On the benefits side, the early estimates allow gaining one year and a half i.e., the first estimate released in September instead of March two years later. The production cycle means that Eurostat can pitch the data to the public three times per year, in March, June and September.

This improved timeliness has been determinant in having new dissemination products built around the data. For instance, in October 2015 Eurostat issued a news release on resource productivity. Another release is scheduled for 2016.

Given the important profile of resource efficiency in the policy agenda, institutional users are very satisfied with this production cycle. The European Commission policymaking services reaffirm their need of more timely indicators, in particular in the context of the European Semester. An approach for early estimates similar to the one described in this paper could be applied to other environmental accounts, with appropriate adjustments to the data availability. Eurostat is reflecting on how the experience of EW-MFA could be extended to other environmental accounts modules, in particular to the module on environmental goods and services sector (EGSS), which includes the important indicator on environmental employment.

Besides the main institutional users, other users also appreciate the improvements. This can be observed from the Eurostat user satisfaction survey, which is an online survey run regularly among users consulting the Eurostat database. After the publication of the early estimates became regular, the percentage of users assessing the timeliness of environment statistics as "very good/good" increased by 7.1% (see green peak in Figure 4). This was the highest increase across statistical domains that year. The increase of users reporting the timeliness was very good/good had a counterpart in the decrease of users reporting the timeliness was adequate.

Figure 4: Users' satisfaction of timeliness, improvement over previous year (%), by statistical domain



Source: Eurostat user satisfaction survey 2014

6. Conclusions

The experience with early estimates in EW-MFA is very positive. Improved timeliness is seen as essential to give environmental accounts, which are relatively new and not sufficiently known, more attention among users. Member States also share this interest and are ready to co-operate as far as they can e.g., reporting early data on voluntary basis in the transmissions to Eurostat. Eurostat intends to develop early estimates for other environmental accounts and the procedures will be enhanced and improved as more experience is gained.

An issue that may raise controversy is the use of modelling in official statistics. One viewpoint is that official statistics should refrain from making use of models because they require assumptions and decisions which affect the output i.e., make the estimates less objective. Accordingly, official statistics should limit themselves to measure the reality, e.g. with surveys or bookkeeping data. In fact, modelling is everywhere as almost any step of statistical production requires some form of model or assumption. A proper use of modelling can allow

official statistics meet the needs of users which otherwise would be unattainable. This is even more relevant for accounting data (national accounts, environmental accounts, etc.) because these operations are based on other official statistics (surveys, censuses, public finance statistics, etc.) and therefore they are less timely. The only way to dramatically improve this time lag is using models. A good example is the quarterly national accounts, which release estimates of GDP and other aggregates a few weeks after the reference period, using models built on top of the annual national accounts (which take several years to be published as final data) and other sources.

In the view of the authors, the point is therefore not whether models should be used in official statistics but how to do it. Requirements are the use of well-established and well-tested models, proper documentation of methods and transparency towards users.