Ecosystem for COllaborative Manufacturing PrOceSses – Intra- and Interfactory Integration and AutomaTION
(Grant Agreement No 723145)

D3.6 Computational Modelling, Simulation and Prediction of Logistics I

Date: 2018-02-27

Version 1.0

Published by the COMPOSITION Consortium

Dissemination Level: Public

Co-funded by the European Union's Horizon 2020 Framework Programme for Research and Innovation under Grant Agreement No 723145

Document control page

Document file: Computational Modelling, Simulation and Prediction of Logistics I
Document version: 1.0
Document owner: CERTH

Work package: WP3 – Manufacturing Modelling and Simulation
Task: T3.3 – Simulation and Forecasting in Production and Logistics
Deliverable type: R

Document status: ☒ Approved by the document owner for internal review
☒ Approved for submission to the EC

Document history:

Version	Author(s)	Date	Summary of changes made
0.1	Thanasis Vafeiadis, Alexandros Nizamis (CERTH)	2018-01-20	ToC and Initial text
0.3	Thanasis Vafeiadis, Alexandros Nizamis (CERTH)	2018-02-06	Contribution on Section 3 and 6
0.4	Nikolaos Alexopoulos, Christos Ntinas (CERTH)	2018-02-09	Contribution on Section 6
0.5	Dimosthenis Ioannidis, Alexandros Nizamis (CERTH)	2018-02-12	Contribution on Sections 4 and 5
0.6	Theofilos Mastos (KLE)	2018-02-14	Contributions to 5.3
0.7	Thanasis Vafeiadis (CERTH)	2018-02-16	Contribution on Section 6
0.8	Alexandros Nizamis, Dimosthenis Ioannidis (CERTH)	2018-02-21	Finalization for the internal review
1.0	Alexandros Nizamis (CERTH)	2018-02-27	Final version submitted to the European Commission

Internal review history:

Reviewed by	Date	Summary of comments
Matteo Pardi (NXW)	2018-02-26	The content of the deliverable is fine. Just a minor comment
Dimitris Gkortzis (ELDIA)	2018-02-26	Approved with no comments

Index:

1 Executive Summary

The present document is a deliverable of the "***Ecosystem for COllaborative Manufacturing PrOceSses – Intra- and Interfactory Integration and AutomaTION***" - (COMPOSITION) project, funded by the European Commission's Directorate - General for Research and Innovation (D-G RTD) under Horizon 2020 Research and Innovation programme (H2020). The deliverable presents an initial version of Computational Modelling, Simulation and Prediction of Logistics developed until M18 of the project.

The COMPOSITION project needs simulation and prediction in both intra- and inter-factory scenarios. This report is focused on inter-factory scenarios which are related to logistics. In this first stage of the project and *Task 3.3 - Simulation and Forecasting in Production and Logistics* the research has been conducted to use cases which were selected as the cases with the highest priority from both pilot and technical partners of the project. This report alongside with *D3.4 - Computational Modelling, Simulation and Prediction of Production I* which is published in parallel with D3.6, present the total work has done in Task 3.3.

The final version of the COMPOSITION's Simulation and Prediction engine and of the corresponding analysis results will be delivered at M28 with the second part of this deliverable, *D3.7 Computational Modelling, Simulation and Prediction of Logistics II*.

2 Abbreviations and Acronyms

Table 1: Abbreviations and acronyms are used in this deliverable

Acronym	Meaning
BMS	Building Management System
DSS	Decision Support System
ERP	Enterprise Resource Planning
GA	Genetic Algorithm
MQTT	Message Queuing Telemetry Transport
IR	Information Retrieval
SSP	Slope Statistic Profile

3 Introduction

3.1 Purpose, context and scope of this deliverable

This document presents the computational modelling, simulation and prediction functions on production developed until M18 of the COMPOSITION project. This document is part of the *Task 3.3 – Simulation and Forecasting in Production and Logistics* mean to design and implement trend analysis techniques for and on significant and key process variables. This deliverable defines the initial approaches for the core set of algorithms, techniques and methodologies dedicated on fill level notification and contractual recyclable material management. With the implementation of such techniques we aim to provide detection of possible optimum pair of routes and weights for waste materials, and a kind of estimations for the bins fullness.

3.2 Content and structure of this deliverable

The content of this deliverable is organized as follows:
Section 4 presents the position of Simulation and Forecasting tool in overall project's architecture. In Section 5 a brief description of the data used for each use case that belongs to the area of computational modelling, simulation and prediction of production, according to *D2.1 – Industrial Use cases for an Integrated Information Management System* is provided. In Section 6 provides a description of the functions and methodologies developed (new) and utilized or modified (existing ones from scientific literature) until M18 of the project, along with their application on the use cases. In Section 7, the next steps of the analysis are briefly described and in Section 8, the conclusions of this initial research are drawn.

4 Simulation and Forecasting tool in Overall COMPOSITION Architecture

This section describes the position of the Simulation and Forecasting Tool in the COMPOSITION project. We provide in this report an overview of the total COMPOSITION architecture and highlighting the Simulation and Forecasting tool's position without giving more descriptions about the components which are connected to it as we do this in more details in *D3.4 Computational Modelling, Simulation and Prediction of Production I* – a report published in parallel with D3.6 and contains the Simulation and Forecasting tool related to intra-factory use cases of the project.

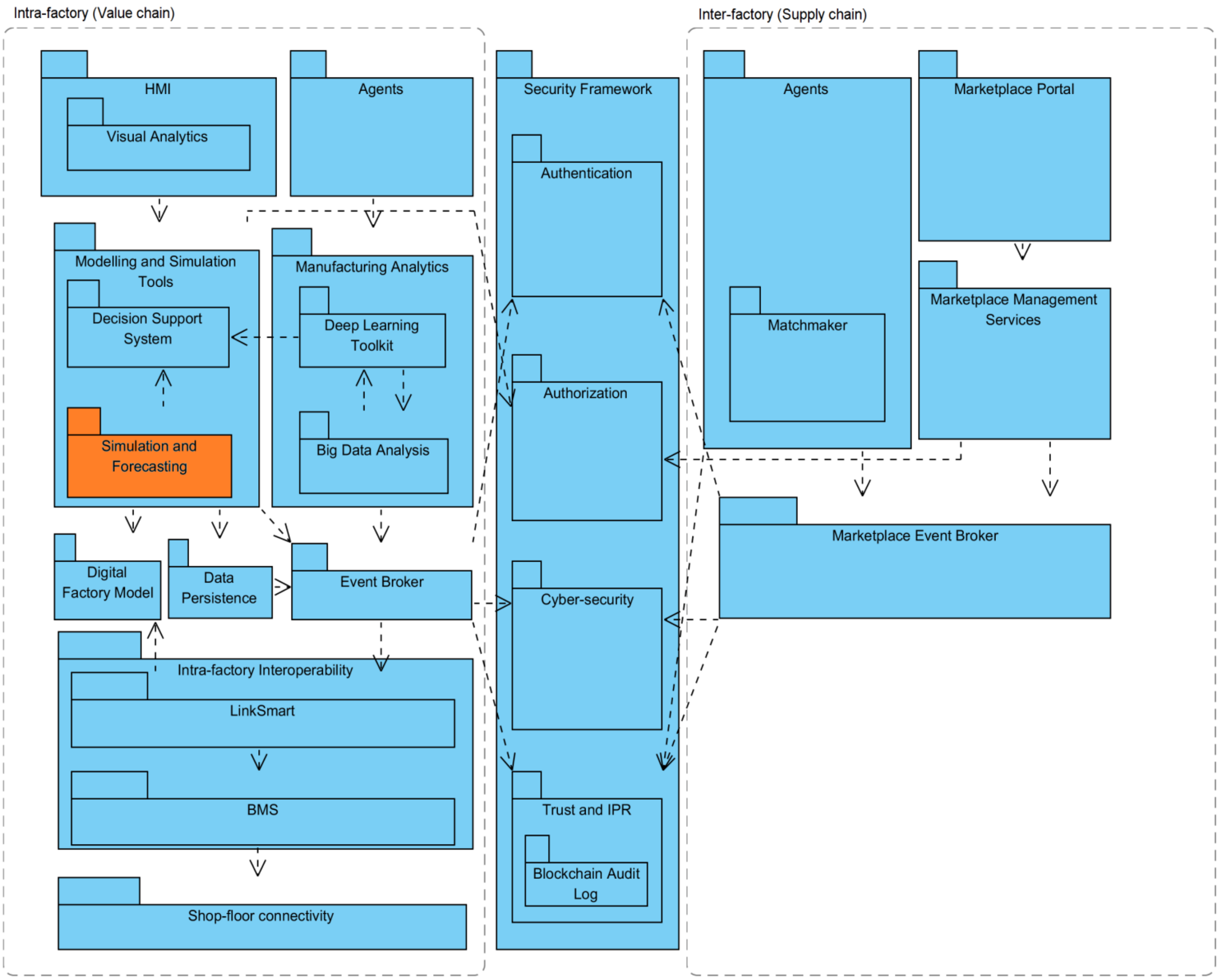

Figure 1: COMPOSITION architecture functional view

As depicted in the Figure 1, the Simulation and Forecasting tool belongs conceptually to intra-factory components. Besides that, as the DSS component is also connected with the inter-factory use cases of the project, it enables the connection of the prediction engine's results with the COMPOSITION Marketplace. The connection of DSS and Simulation and Forecasting tool is presented in D3.4 as it mentioned before.

5 Industrial Data Description

In this section a brief description of the data used for each use case that belongs to the area of computational modelling, simulation and prediction of production, according to *D2.1 – Industrial Use cases for an Integrated Information Management System* is provided.

5.1 Use Cases Prioritisation

Before we started with use cases' data description in this section and the data processing in the following section, it is worth to mention that the work has been done at Task 3.3 by M18 was mainly focused in scenarios which selected as of highest priority from both technical and pilot partners.

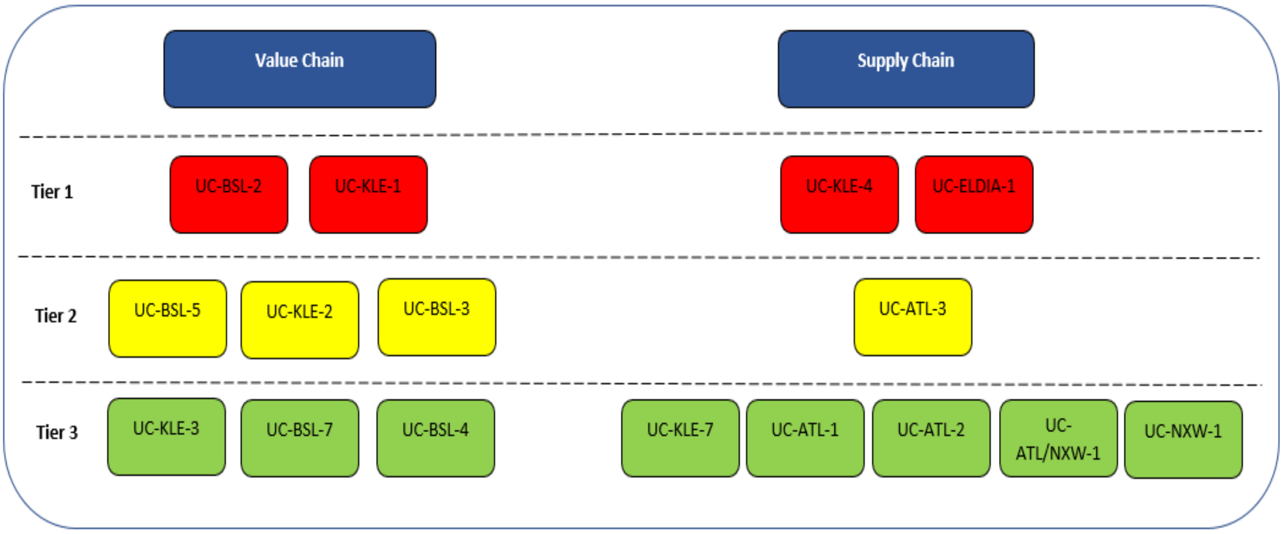

Figure 2: COMPOSITION Use cases prioritisation

As depicted in the figure 3 the UC – KLE 4 Scrap Metal Collection Process and UC – ELDIA 1 Fill Level Notification – Contractual Recyclable Material Management are the highest priority inter-factory scenarios cases. The main focus of this report is the documentation of research and development related to these two use cases.

5.2 UC – ELDIA 1 Fill Level Notification – Contractual recyclable Material Management

The primary goal of this use case is to be able to receive notifications of the fill level of various containers installed at our customers' facilities, thus facilitating the logistics service and improving the reaction time.

ELDIA ERP maintains a database including information about the date of pick up, type, weight and prices of various recyclable materials. This information become available to the project's technical and research partners for further analysis in order to enable possible estimations of the fill level of various containers.

5.3 UC – KLE 4 Scrap Metal Collection Process

The goal of this use case is to optimize scrap metal collection and bidding process by getting better scrap metal prices, minimizing costs and receiving a fast and efficient service. A dataset is generated by the company's ERP system. The ERP maintains a database including information about the produced scrap metal and the price of it. This set of data is extracted from ERP in excel as a report file. In 2016 around 1.000 tons of scrap metal were produced with an average price of € 120 - 150/ ton.

6 Data Preparation and Processing

This section provides a description of the functions and methodologies developed (new) and utilized or modified (existing ones from scientific literature) until M18 of the project, along with their application on the use cases.

6.1 Fill Level Sensor Deployment

For the COMPOSITION use cases related to bins fill level notification (UC-KLE 4 and UC-ELDIA 1); a fill level monitoring sensor has been developed. This sensor is able to provide data to the Simulation and Prediction tool in order to enable the estimation of the date in which the bin will be full.

Figure 3: Fill level sensor's position

For the fill sensor data the HRL-MaxSonar-EZ MB1013 Ultrasonic Sensor was used to capture raw measurements of distance of the waste heap from the deployment point. The sensor has its own filtering mechanisms to provide reliable measurements and provides the result via PWM output. It has to be noted that the sensor will be deployed with custom protection from rain. In case that the custom protection is not sufficient, the IP67 counterpart of the sensor will be used, HRXL-MaxSonar-WRT MB7380, with minor changes on the system. The sensor and the communication is controlled by the STM32L053c8t6 low power microcontroller. The microcontroller also checks for faulty measurements and repeats the measurement process again and if the measurement is faulty sends an error flag. The wireless communication is carried out via the sx1272mbas LoRa Module for STM. The system is powered by 4 AA batteries.

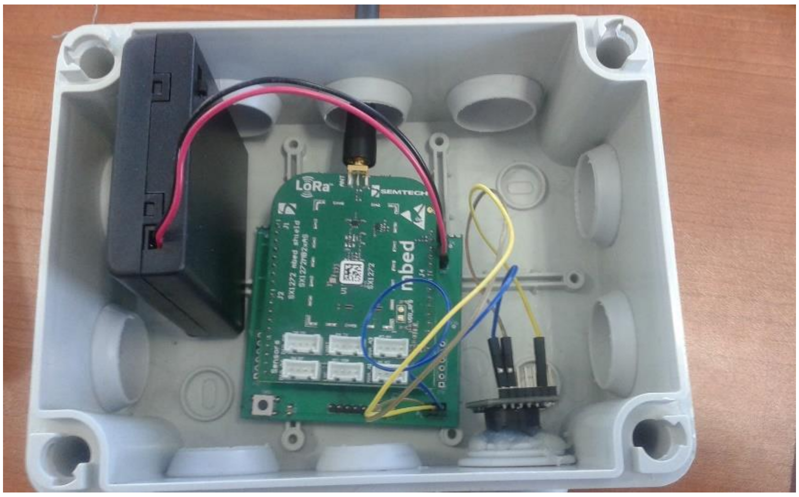

Figure 4: Fill level sensor system overview

Data are transferred via the LoRa low power protocol to the LoRa Gateway. The gateway to be used is the LoRank 8. It has to be noted that LoRa protocol allows transmission of only very small packets of data, so this means that only raw measurements of distance are transferred along with the measurement of battery level and id of bins. The gateway that is connected to the internet via Ethernet or Wi-Fi publishes those data on an MQTT topic on the cloud and a listener, connected to the same broker as the gateway, reprocesses the data to convert the raw distance measurement into a fill percentage and derive the json object format to be sent on the destination platform. The Building Management System (BMS) component of the project will be the destination platform.

6.1.1 Slope Statistic Profile (SSP)

For the analysis of fill level sensor data, we will apply the Slope Statistic Profile methodology. Slope Statistic Profile (Vafeiadis, 2011), denoted hereafter as SSP, is a method that detects the single structural break T, denoted hereafter as incident, in a time series using a standard parametric linear trend test, denoted hereafter as t-statistic. The t-statistic is calculated on overlapping sliding data windows of size w with sliding step one, along the time series. By this way we obtain the profile of the t-statistic, denoted as $\{\widetilde{U}_i\}$, for $i = 1 + [\frac{w}{2}], \dots, n - [\frac{w}{2}]$, where $[x]$ is the integer part of x. The form of this profile depends on time series characteristics, i.e. the strength of the autocorrelation, the distribution of the residuals, the strength of the linear trend, as well as, the size of the sliding data window w. The profile of the t-statistic $\{\widetilde{U}_i\}$, exhibits small fluctuations (glitches) due to edge effects of the local data windows and therefore the profile curve is smoothed using a zero-phase filter of a small order, set to about 5% of w. Such a small filter order removes the glitches in $\{\widetilde{U}_i\}$, but maintains its original signature. The smoothed value of $\widetilde{U}_i$ is denoted as U_i and referred to as U-profile. In a short presentation of the method below, we will assume the situation from no trend to a positive trend. Other types of change between no trend and trend can be treated similarly.

The t-statistic for the parametric linear trend test is $t = \hat{\beta} / s(\hat{\beta}) \sim t_{w-2}$, where $\hat{\beta}$ the trend parameters and $s(\hat{\beta})$ is the standard error of $\hat{\beta}$. The null hypothesis of no trend is rejected at the significance level a if $|t| \geq t_{w-2,1-a/2}$.

SSP methodology will be applied on time series of fill level so as to provide notifications and warnings regarding the fill level of the recycle bins. SSP methodology hasn't tested yet on use case data. The deployment of the fill level sensor and the research are ongoing and the results of the application of the method will be presented on the next version of this document.

6.2 UC – ELDIA 1 Fill Level Notification – Contractual recyclable Material Management

6.2.1 Methodologies

6.2.1.1 Genetic Algorithms

People use search engines to find information they desire with the aim that their information needs will be met. Information retrieval (IR) is a field that is concerned primarily with the searching and retrieving of information in the documents and also searching the search engine, online databases, and Internet. Genetic algorithms (GAs) are robust, efficient, and optimized methods in a wide area of search problems motivated by Darwin's principles of natural selection and survival of the fittest.

IR system searches for the matches in the document databases and, thus, retrieves search results of the matching process. However, based on the relevance, the user will then evaluate and display the search results. The relevance of the document is very important to the user. If the user feels that it is a relevant document, he finishes the search while else user continues to search in the document database by reformulating the query until the relevant documents that will satisfy users' information needs are retrieved.

GA is a probabilistic algorithm simulating the process of natural selection of living organisms and finally coming up with an approximate solution to a problem (Holland, 1975), (DeJong, 1975) & (Goldberg, 1989). In GA implementation, the search space is composed of candidate solutions (called individuals or creatures) to an optimization problem to evolve better solutions; each represented by a string is termed chromosome. Each chromosome has an objective function value, called fitness. A set of chromosomes together with their associated fitness is called a population. This population, at a given iteration of the genetic algorithm, is called a generation. In each generation, the fitness of every individual in the population is evaluated from the current population based on their fitness value and modified to form a new population. The new population is then used in the next iteration of the algorithm.

GA terminates when either a maximum number of generations has been produced or a satisfactory fitness level has been reached for the population. If the algorithm has terminated due to a maximum number of generations, a satisfactory solution may or may not have been reached. The working of the genetic algorithm depends upon the constraint of how well we choose our initial random keywords.

6.2.2 Applications

In the current contractual recyclable material management dataset (see Table 2) there were 4 recyclable materials (waste = 1, wood = 2, plastic =3 and paper = 4) recorded for the years of 2015 and 2016. In order to apply the GA we have chosen waste and wood materials, which are recorded for all customers for both years. The fitness function is the ratio or routes and weights that need to be minimized (the company needs to carry as much weights as possible with as few as possible routes. Thus, the target value of this fitness function is the minimum ratio of all customers for a specified material. In Table 3, one can see the average values of fitness function for all waste and wood materials, for all customers and for the years 2015 and 2016. The Objective bounds of routes and weights per material are given in Table 4. The ratios of routes per weights for all customers, materials and years are also given in Figure 5: Ratios of routes per weights for all customers, materials and years. As input in the GA are the bounds of routes and weights along with the minimum ratio of all customers per material and per year (see the highlighted values in Table 3).

Table 2: Contractual recyclable material management dataset

A.A.	Customer	Description	January	February	March	April	May	June	July	August	September	October	November	December	Sum
1	Customer A 2015	Route_Material_1	2	4	6	6	6	8	8	5	5	7	6	6	69
		Weight_Material_1	10.070	17.960	35.420	27.070	23.840	46.860	44.040	25.190	28.240	35.300	23.830	27.520	345.340
		Route_Material_2	2	3	3	4	5	6	3	4	4	3	5	4	46
		Weight_Material_2	0.750	1.230	1.380	1.920	2.020	2.610	1.080	2.030	1.880	1.390	2.130	2.840	21.260
	Customer A 2016	Route_Material_1	3	6	7	8	8	6	6	5	8	7	6	3	73
		Weight_Material_1	14.200	29.420	36.400	40.700	39.940	32.360	33.050	23.570	41.510	34.500	24.230	13.700	363.580
		Route_Material_2	4	4	5	5	6	4	3	4	4	4	4	2	49
		Weight_Material_2	3.440	1.960	3.990	1.870	4.630	1.480	1.270	2.360	1.950	1.780	1.630	0.700	27.060
2	Customer B 2015	Route_Material_1	1	1	2	3	2	3	1	2	3	3	1	2	24
		Weight_Material_1	1.490	2.780	5.110	6.220	3.410	6.220	2.790	2.960	5.710	6.420	0.900	6.540	50.550
	Customer B 2016	Route_Material_1	1	1	1	3	1	2	2	1	2	1	1	2	18
		Weight_Material_1	3.030	3.810	3.050	6.350	2.220	5.190	4.180	2.330	5.560	2.670	2.940	5.460	46.790
3	Customer Γ 2015	Route_Material_1	6	11	10	10	10	11	10	6	8	13	10	12	117
		Weight_Material_1	20.540	40.210	27.970	29.750	30.790	32.310	29.440	24.110	25.000	64.700	27.390	29.950	382.160
		Route_Material_2	3	3	3	3	3	4	4	2	3	4	4	5	41
		Weight_Material_2	1.030	1.190	1.290	1.750	1.230	1.870	1.610	0.710	0.950	1.670	2.400	1.930	17.630
		Route_Material_3	2	2	2	1	2	2	3	-	1	1	1	1	18
		Weight_Material_3	0.570	0.200	0.270	0.190	0.320	0.190	0.320	-	0.190	0.190	0.190	0.040	2.670
		Route_Material_4	4	5	4	4	5	5	6	2	6	5	5	6	57
		Weight_Material_4	2.570	2.200	2.690	1.940	3.040	2.090	2.770	1.310	3.400	3.660	2.410	2.660	30.740
	Customer Γ 2016	Route_Material_1	8	12	12	14	10	10	12	8	14	12	14	16	142
		Weight_Material_1	25.920	37.010	32.250	37.960	26.190	35.970	36.830	29.740	43.640	33.330	40.010	53.420	432.270
		Route_Material_2	3	2	7	5	5	4	5	4	6	7	7	5	60
		Weight_Material_2	1.120	0.940	2.770	2.000	3.790	4.290	1.980	5.820	2.530	2.950	4.950	1.930	35.070
		Route_Material_3	3	-	-	1	1	-	-	-	1	-	-	2	8
		Weight_Material_3	0.380	-	-	0.250	0.210	-	-	-	0.180	-	-	0.620	1.640
		Route_Material_4	4	6	7	7	5	6	6	4	6	7	6	6	70
		Weight_Material_4	2.040	4.020	3.850	3.100	2.460	4.000	2.930	1.780	2.970	3.750	3.560	3.660	38.120
4	Customer Δ 2015	Route_Material_1	10	10	11	11	13	12	14	8	12	17	13	10	141
		Weight_Material_1	31.690	32.180	39.430	41.140	44.510	36.170	40.550	25.620	36.340	38.820	49.940	28.110	141.000
	Customer Δ 2015	Route_Material_1	9	10	10	11	11	13	15	7	14	11	12	10	133
		Weight_Material_1	26.330	37.590	40.800	34.720	34.450	37.990	45.430	20.380	43.080	32.140	36.890	37.550	427.350

Table 3: Fitness function average value for material / customer for years 2015 and 2016.

Years	**2015**		**2016**	
Materials	**Waste**	**Wood**	**Waste**	**Wood**
Customer A	**0.17**	**1.41**	**0.18**	1.16
Customer B	0.30	1.66	0.26	**0.68**
Customer C	0.20	-	0.27	-
Customer D	0.26	-	0.25	-

Table 4: Objective bounds of routes and weights per material

Years	**2015**				**2016**			
Materials	**Waste**		**Wood**		**Waste**		**Wood**	
	Routes	**Weights**	**Routes**	**Weights**	**Routes**	**Weights**	**Routes**	**Weights**
Min	1	1	2	0.71	7	20.38	2	0.94
Max	17	65	5	2.4	15	45.43	7	5.82

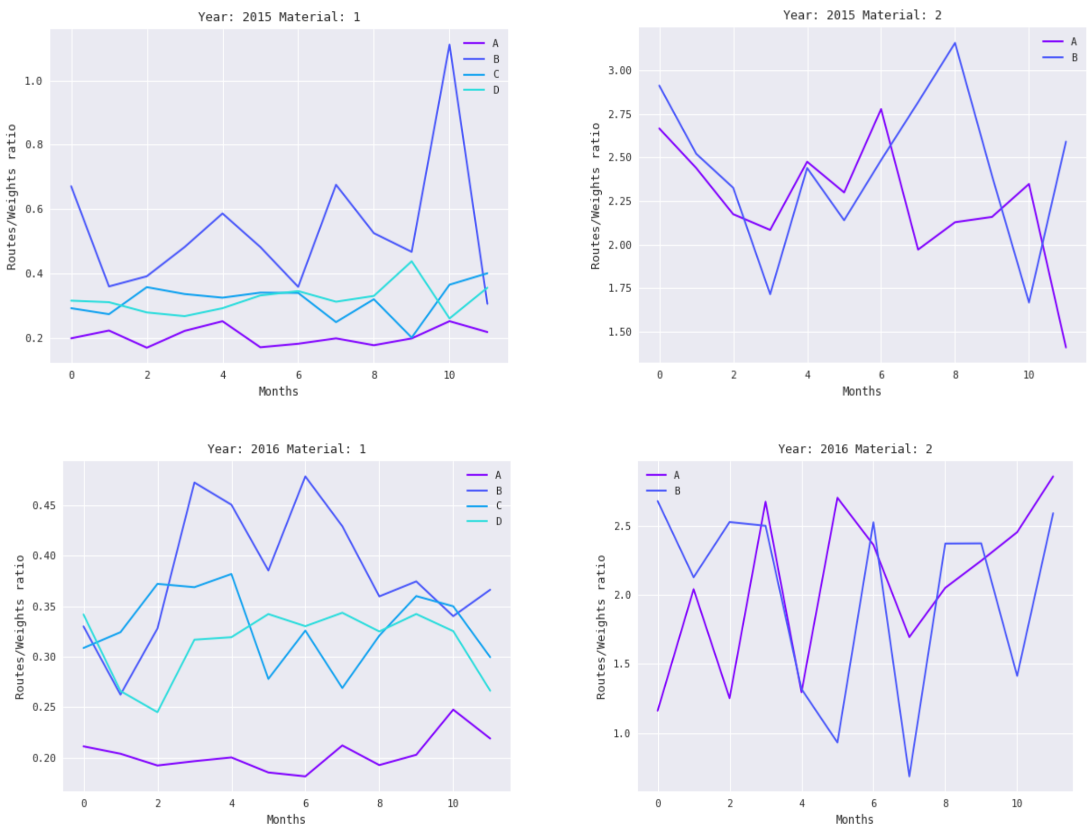

Figure 5: Ratios of routes per weights for all customers, materials and years

Thus, the results from the application of GA on the various materials (waste and wood) and for years 2015 and 2016 are given in Table 5. For the year 2015 the optimum pair of routes and weights for waste material is [9, 53] and wood material the optimum pair of routes and weights is [4, 3]. For the year 2016 the optimum pair of routes and weights for waste material is [7, 39] and wood material the optimum pair of routes and weights is [2, 3]. The optimum pair denotes what would be the number of routes and the number of carried weights so as the company to have no losses.

Table 5: GA selections

Years	2015				2016			
Materials	Waste		Wood		Waste		Wood	
	Routes	Weights	Routes	Weights	Routes	Weights	Routes	Weights
GA	9	53	4	3	7	39	2	3

6.3 UC – KLE 4 Scrap Metal Collection Process

This use case starts from the fill level monitoring using the aforementioned fill level sensor's measurements in KLEEMANN factory's scrap metal bin. After that, the action is transferred out of the factory. The KLEEMANN's Marketplace agent generates an automated bidding process. As soon as, the Matchmaker finds the supplier agents who offer waste management services, these agents are informed by the system in order to send their requests. By the time that the bidding process is expired, the Matchmaker evaluates the offers and suggests the best one to the KLEEMANN's requester agent. In the next step of the use case the purchase manager accepts the suggested offer by the system or he selects one examining the rest available offers. Finally the selected agent/company is informed by the system and the scrap metal collection process is completed as soon as the scrap metal is collected by the waste management company.

The involvement of the simulation and forecasting tool in this use case is related to the prediction of the bin's fill level in order to generate earlier the bidding process and enable the faster completion of the whole process.

The Slope Statistic Profile methodology which is presented in 6.1.1 section has been selected to be applied to the fill level sensor's data. More precisely, the SSP will be applied on time series of fill level in order to provide notifications regarding the fill level of the recycle bin.

As the permanent deployment of the fill level sensor in KLEEMANN's bin is an ongoing activity, the applied SSP methodology's results will be presented in the second iteration of this report coming at M28.

7 Next Steps

The future work at *Task 3.3 Simulation and Forecasting in Production and Logistics* will be mainly focused at procedures related to:

- For the UC-ELDIA-1 Fill Level Notification – Contractual recyclable Material Management the data coming from the installed fill level sensor will be used by the prediction engine. Moreover further research and development will be conducted based on ELDIA new data
- Further research and development will be conducted for UC – KLE 4 Scrap Metal Collection Process based on the data coming from the installed fill level sensor. The SSP method will be tested and maybe other approached will be examined.
- Dockerization of the future developed algorithms in order the prediction results to be available to DSS and the COMPOSITION Marketplace as well

The results of the previous mentioned procedures of Task 3.3 will be reflected at the next and final version of this report. This work will be presented in *D3.7 Computational Modelling, Simulation and Prediction in Logistics II in M28.*

8 Conclusions

In conclusion, this deliverable describes the effort spent from M4 to M18 and represents the current status of *T3.3 – Simulation and Forecasting in Production and Logistics* of WP3. More precisely, this report documents the results in the part of logistics as the production results are documented in D3.4 - Computational Modelling, Simulation and Prediction of Production I which will be submitted in parallel with the current report. The complete work of Task 3.3 related to logistics will be presented in *D3.7 Computational Modelling, Simulation and Prediction in Logistics II* in M28.

As it is perceived in this document, in the first steps of Task 3.3 the research and development were focused on the architecture design, the analysis of existing pilot partners' data, the selection of the proper methodologies for this data, and the development of the aforementioned methodologies for the use cases with the highest priority.

Based on the analysis of the applied methods, the Genetic Algorithm for ELDIA contractual recyclable material management indicates the optimum pair of routes and weights for waste and wood materials in order to enhance the decision making by the ELDIA purchase manager. Furthermore, the application of the Slope Statistic Profile methodology will enable the COMPOSITION system to offer a kind of predictions for the bins fullness. By using this algorithm's estimations, the scrap metal bidding process will be able to start and be completed earlier. Besides this, the system will be able to offer to ELDIA not only a notification of the fill level of the containers but also an estimation of the future container's fullness in order to provide a more complete solution for the decision making.

9 List of Figures and Tables

9.1 Figures

9.2 Tables

10 References

(DeJong, 1975) K. A. DeJong, An Analysis of the Behaviour of a Class of Genetic Adaptive Systems, University of Michigan, 1975.

(Goldberg, 1989) D. E. Goldberg, Genetic Algorithms in Search, Optimization, Machine Learning, Addison Wesley, 1989.

(Holland, 1975) J. H. Holland, Adaptation in Natural and Artificial Systems, The University of Michigan Press, Ann Arbor, Mich, USA, 1975.

(Vafeiadis, 2011) Vafeiadis, T., Bora-Senta, E., Kugiumtzis, D., (2011). 'Estimation of linear trend onset in time series'. Simulation modelling – Practice and Theory, Vol. 19, Issue 5, p 1384 - 1398.