Abstract—A linear programming (LP) approach is proposed to estimate the maximum flow in the European air traffic network, based on empirical traffic data and capacity constraints at airports and in Area Control Centres. In order to validate the model and the associated results, we treat historical data collected on 1 July 2012 as the empirical maximum flow and compare them with the theoretical maximum established by the LP model. The results show that the proposed model, despite its relatively simple structure and assumptions, has captured the overall congestion status and the saturated state of the network, which is verified statistically. Meanwhile, we utilize data on Air Traffic Flow Management (ATFM) delays to validate a queuing model by comparing the empirical data with predicted data obtained from Little’s theorem. The ATFM delays, when combined with the proposed LP model, suggest a viable way of identifying and quantifying several capacity factors that impact on network capacity such as airspace availability and ATFM. Future work will describe and predict more accurately the air traffic network in Europe.

Keywords-capacity factor, linear programming, maximum flow, network capacity, SESAR.

I. INTRODUCTION

In Europe, air traffic has doubled in the past 20 years and an average annual growth of 0.6 % has been predicted from 2013 to 2019 [1]. Although traffic growth has flattened and the performance of the European air traffic network has improved, the congestion at busy airports and in Area Control Centres (ACCs) still remains severe. The network includes the member States of the European Union and the European Organization for the Safety of Air Navigation (EUROCONTROL). In order to cope with this congestion, the Single European Sky (SES) Air Traffic Management (ATM) Research (SESAR) program was launched to transform European ATM from an airspace-based to a trajectory-based system. Trajectory-based operation is at the heart of the Concept of Operations (ConOps) which enables the European ATM system to be considered as a continuum[2, 3] and consequently network capacity is defined as an essential key performance area of SESAR.

In general, capacity is defined as the maximum number of aircraft that can be handled safely over a period of time [4]. Numerous studies have presented different models to quantify capacity at airports and in en-route airspace [5-17] and in particular on bottlenecks of the runway system and air traffic controller workload respectively [9, 15, 18].

Trajectory-based operations require capacity to be modelled and quantified at a network level. EUROCONTROL has defined network capacity to be the network throughput taking traffic demand patterns and the network effect of airports and airspace into account [19]. However, this definition does not capture the influence of capacity factors introduced by the new ConOps.

Currently, Air Traffic Flow Management (ATFM) delay is used as a key performance indicator to monitor network capacity [20]. However, ATFM delay is not a direct measure of capacity but a proxy that reflects the extra time caused by capacity shortages at airports and in en route airspace. As it is not a direct measure, there is an inherent level of inaccuracy in the use of delay that makes it difficult for stakeholders of the ATM system to measure performance and to make key operational decisions.

The Future ATM Profile, which comprises a set of capacity estimation methods, is currently used to derive the capacity requirements from network to ACC level. As a result, every air navigation service provider is required to build a Local Convergence & Implementation Plan to improve the capacity under their responsibility. However, the relationship between the ACC capacities determined in this way and network capacity is unknown.

To date there has been limited research on direct methods for the modelling and estimation of air traffic network capacity. The only approach available in the public domain was presented by Donohue [21] referred to as the Macroscopic Air Transportation Capacity Model (MCM). In this approach, using data from the USA, airports are assumed as bottlenecks within the US ATM network and the inefficiencies in airspace can be ignored. Therefore, in this model, the network capacity is equal to the sum of airport capacities.

However, when latter applied to European airspace, it was found that the inefficiencies in the airspace could not be
ignored, with the implication that the original assumptions are not transferrable across airspaces [22]. Lulli and Odoni [23] also showed that the inefficiencies in European airspace have increased the complexity of ATFM. Clearly, network capacity determination based on the summation of airport capacities has significant limitations. Therefore, a network capacity estimation method that accounts for all relevant factors is required, and is the subject of this paper.

Traditionally, a transport network is composed by a set of nodes and links. Each link can only handle a limited amount of traffic that is its capacity. The maximum flow in a network is the theoretical maximum amount of traffic that can be handled by the network [23, 24]. The actual traffic flows must be lower than the theoretical maximum for compliance with capacity constraints.

As noted previously, EUROCONTROL’s definition of network capacity does not capture the influences of all factors that affect capacity i.e. capacity factors. We argue that the gap between theoretical and empirical maximum of network flow can be explained by the inefficiencies in the capacity factors. To support this argument, this paper proposes a linear programming (LP) approach that not only estimates the maximum flow in the European air traffic network, but also identifies those capacity factors mentioned above via regression analysis. Specific contributions and/or potential impact made in this paper include the following.

- The proposed LP model captures the overall saturated state of the network and identifies bottlenecks, which is consistent with empirical observations.
- A well-defined linear relationship between the theoretical and empirical maximum flows is uncovered by regression analysis. This finding provides the first clue on the quantified effect of the capacity factors.
- The proposed LP model, when combined with the ATFM delays, holds promise in further identifying each individual capacity factor and quantifying its influence on network capacity. This will be done in future research.

The rest of this paper is organized as follows. Section II provides general background of the problem of interest. Section III discusses in detail methodology employed for estimating maximum network flow. Findings and accompanying discussions are presented in Section IV and Section V. Based on these findings, Section VI proposes some future research directions. Some concluding remarks are provided in Section VII.

II. BACKGROUND

This section introduces background materials relevant to the European air traffic network, network capacity, maximum flow and capacity factors.

A. European Air Traffic Network

In the European air traffic network displayed as a graph, the nodes represent airports and ACCs. A critical notion is connectivity, which can be defined as a binary state that exists between any two nodes in the network, and takes value 1 if the two nodes are connected by a link and 0 otherwise. Unlike many traditional transport networks, the capacity constraints in an air traffic network are imposed on the nodes (airports and ACCs) rather than on the links.

The declared capacities at airports and in en-route airspace have been used to prevent the airspace from overloading by applying the mechanisms of ATFM, which include re-routing, altitude change and the imposition of regulations. Air traffic at European airports and in en-route airspace is required to comply with the declared capacities.

Based on these characteristics, the European air traffic network can be regarded as a capacitated transport network and, as a result, the traffic flow through it cannot exceed the theoretical maximum.

B. Network Capacity

As argued previously, network capacity ($C_{net}$) is equal to the maximum flow ($F_{max}$) reduced by a series of degradation-parameters ($D$). Each such parameter represents the inefficient performance of its correlated capacity factor and parameters take a value between zero and one. If the performance of a given capacity factor do not influence traffic flows in a network, the correlated degradation-parameter equals to one.

$$C_{net} = F_{max} - D_1D_2D_3...$$

The maximum flow and capacity factors are introduced in next subsections.

C. Maximum Flow

In graph theory, the maximum flow is determined by network topology, capacity constraints on nodes and links and traffic patterns. The network topology is the geographic characteristics of connectivity between nodes. The maximum flow can be formulated as a function of network topology ($T_{net}$), node capacities ($C_{airport}$ and $C_{airspace}$) and traffic demand patterns ($P_d$). It is formulated as:

$$F_{max} = f(T_{net}, C_{airport}, C_{airspace}, P_d)$$

D. Capacity Factors

In the case of the European air traffic network, capacity factors can be categorized into the following three groups: airport, airspace and network. Airport and airspace capacity factors not only influence the capacity locally but also consequently influence the maximum flow at network level. Taxonomies of airspace and airport capacity factors have been developed by [16, 25].

In addition to the airport and airspace capacity factors, network capacity factors influence capacity directly at a macroscopic level. However, little research exists regarding the network capacity factors, and this is an important area for future research.
III. METHODOLOGY

The methodology of estimating the maximum flow in the European air traffic network is depicted in Fig. 1 and involves three steps: definition, model building and validation. The following subsections show the characteristics of these steps.

A. Definition

1) Metrics selection: Due to the fact that the terminologies in graph theory and transport network analysis are not consistent, it is necessary to unify some useful terminologies and metrics.

a) Theoretical Maximum Flow (TMF): It is the theoretical upper bound of the traffic flow in a capacitated network. It is based on network topology and capacity constraints. The definitions of TMF at airports, in ACCs and network are:
   - TMF at an airport: the maximum number of take-offs and landings that can be served by a given airport in a period of time.
   - TMF in an ACC: the maximum number of flights that served by a given ACC in a period of time.
   - TMF in a network: the maximum number of flights that can be handled by the network in a period of time.

b) Empirical Maximum Flow (EMF): This is the maximum traffic flow that can be realistically handled by a network. It is also the network capacity that can be practically delivered. The metrics of the EMF are in line with the TMF. In practice, congestion is a reflection of the shortage of capacity in the transport system and consequently the traffic on congested days can be assumed to be the EMF. The indicator of congestion is the average ATFM delay per flight.

c) Utilization rate ($\rho$): This is the ratio of the actual traffic to the capacity and reflects the level of congestion. The higher the utilization rate, the lower the available capacity remaining.

d) Lagrange multiplier ($\lambda$): This is the sensitivity of the optimized quantity with respect to the change in capacity constraints. Lagrange multipliers are zeros on any non-bottleneck node. This is also the marginal benefit of increasing capacity on the correlated bottleneck and the higher the value of $\lambda$, the greater the contribution to network capacity from the bottleneck.

2) Data Processing:

a) Flight profiles:

As noted previously, the air traffic on a congested day can be used as the EMF, a dataset of air flight profiles on 1 July 2012 was chosen. The average ATFM delay per flight on this day was the highest in 2012 except three other days with industrial actions. In addition to the high ATFM delay, this day is in the European summer which is the season with the highest traffic demand in a year.

The total scheduled flights were 28,904 and the flight profiles of 28,885 flights were recorded by radar. In total, 28,776 flight profiles are used and the other 109 flights are ignored. The ignored flights were either using unrecognized airports or passing unrecognized airspaces.

These flights are categorized into three categories namely intra-European (IntraEU) flight, intercontinental (InterCon) flight and overflight (see Table I).

<table>
<thead>
<tr>
<th>Flight Category</th>
<th>IntraEU</th>
<th>InterCon</th>
<th>Overflight</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Network Flight</td>
<td>22,578</td>
<td>5,853</td>
<td>454</td>
<td>28,776</td>
</tr>
<tr>
<td>Airport Operation</td>
<td>45,156</td>
<td>5,853</td>
<td>0</td>
<td>51,009</td>
</tr>
<tr>
<td>ACC Flight</td>
<td>81,469</td>
<td>26,979</td>
<td>1,482</td>
<td>109,300</td>
</tr>
</tbody>
</table>

The airport operations and ACC flights used by the flights in different categories are listed in Table I. It is clear that each flight in different categories introduces different numbers of operation to airports and ACCs. For example, an IntraEU flight requires two airport operations and an InterCon flight only requires one. The quantity of ACC flights is determined by the length of the flight path.

b) Capacities:

The capacities used in this chapter are airport declared capacities and ACC capacity baselines. The capacity data used in this research are either published in [26, 27] or recorded in the dataset of flight profiles. Although there are inherent inaccuracies in these published capacities, we use fixed figures of capacities to calculate the maximum flow and the inaccuracies will be assessed in future research.

c) ATFM delays:

Donohue [21] suggested that Little’s theorem can be used to validate capacities. In this paper, we use Little’s theorem to compare the actual queue lengths to the theoretical predictions. The actual queue lengths are calculated by multiplying ATFM delays and capacities together. The data of ATFM delays on 1
July 2012 were gathered from the public website of EUROCONTROL [28].

The causes of ATFM delays used by EUROCONTROL are divided into 15 categories [29]. Each category can be mapped to a correlated capacity factor. Quantification of these causes of ATFM delays enables us to calculate the degradation-parameters.

B. Model building

1) Network Topology

This subsection introduces the constituent ACCs and airports in the European air traffic network.

a) ACCs: According to [26], European airspace is controlled by 64 ACCs.

b) Airports: Based on the data of flight profiles, 784 airports are identified. For the sake of mathematic tractability, it is necessary to reduce the number of nodes. In order to keep the characteristics of congestion and geographic conditions of airports, the nodes are categorized into the following two groups:

- Busy airports: Based on [27, 30], 67 busy airports are chosen to observe the phenomenon of congestion on bottlenecks. The flight operations at these 67 airports represent 62% of total flights on the day.
- Aggregated airports: Given that the traffic at other 717 less-busy airports is only 38% of total traffic, it is reasonable to focus at a macroscopic level rather than at an individual less-busy airport. The remaining 717 airports are merged into 64 aggregated nodes. The airports connecting to the same ACC are merged into an aggregated node. Merging airports in line with the geographic connection not only reduces the scale of the network but also maintains the characteristics of connectivity.

c) Connectivity: A connectivity matrix represents the relation between adjacent nodes and is usually used to depict the network topology. It is also known as an adjacency matrix. This network contains 197 nodes including 67 busy airports, 64 aggregated airports, 64 ACCs and two imaginary ACCs which representing all external airspaces.

2) Capacity Constraints: The airport capacities and ACC capacity baselines are used as capacity constraints. Although the capacity data of less-busy airports are unavailability in public, we assume that the capacity of each aggregated node equals to 40% of the adjacent ACC.

Due to the fact that all flight profiles on the day are used, the capacities of airports and ACCs are converted from an hourly basis to a daily basis by multiplying 16 and 24 respectively.

3) Traffic Demand Patterns: These are the patterns of the amount of flight operations in a given time period. Such patterns are related to a given area, route, location or service. The data of flight profiles effectively reflects the pattern of air traffic demand on 1 July 2012.

4) Maximum Flow Estimation

Consider a capacitated network \( G = (P, W, C) \) with a nonnegative capacity \( C \) associated with each constituent node. \( P \) and \( W \) represent the information of flight paths and the set of origin-destination (O-D) pairs respectively.

\( \forall (i, j) \in W, \text{ let } P_{ij} \text{ to be the set of flight paths connecting airport } i \text{ to airport } j. \forall p \in P, p=\{i, k_1, k_2, ....., k_m, j\}, \text{ where } i, j \text{ are airport nodes and } k_1, k_2, ....., k_m \text{ are ACC nodes. The flow on the path } f_p \text{ is nonnegative.} \)

A path-node matrix \( \delta_{pv} \) that contains all paths \( (p) \) and nodes \( (v) \) is built:

\[
\delta_{pv} = \begin{cases} 
1 & \text{if } v \in p \\
0 & \text{otherwise}
\end{cases}
\]  

Our maximum flow problem can be stated as follows: \( (A \) denotes the set of paths using airports)

Object function:

\[
\max \sum_{i \in A} f_i
\]  

Subject to capacity constraints:

\( \forall i \in \text{Airports} \) (the set of airports)

\[
\sum_{p \in P} \delta_{pi} f_p \leq C_i
\]  

\( \forall k \in \text{ACCs} \) (the set of Area Control Centres)

\[
\sum_{p \in P} \delta_{pk} f_p \leq C_k
\]  

Fig. 2 depicts the topology of the European air traffic network which comprises 197 numbered nodes. The flight path \( f_1 \) and \( f_2 \) stand for intraEU flights. The flight path \( f_3 \) and \( f_4 \) are interCon flights, and \( f_5 \) is overflight.
In order to reflect the uneven distribution of air traffic in Europe, the objective of our model is to maximize the traffic flow at the 67 busy airports. The model can be stated as follows.

Objective function: \( \text{BA} \) denotes the set of paths using busy airports

\[
\max \sum_{i \in \text{BA}} f_i
\]  

Subject to

\[
\begin{bmatrix}
F_1 \\
F_2 \\
\vdots \\
F_{28776}
\end{bmatrix}
\begin{bmatrix}
f_1 \\
f_2 \\
\vdots \\
f_{28776}
\end{bmatrix}
\leq
\begin{bmatrix}
C_1 \\
C_2 \\
\vdots \\
C_{197}
\end{bmatrix}
\]

C. Model Verification

The LP model contains three sub-models: flight path identification, network topology identification and maximum flow estimation. The methods used for these sub-models are introduced below.

1) Flight Path Identification: Each flight profile in the dataset contains origin, destination airports and a group of ACCs. External airports and ACCs are screened out and a flight path matrix is built in accordance with the 28,776 flights.

2) Network Topology Identification

a) Connectivity matrix: Based on the flight path matrix, a connectivity matrix is created by searching the adjacent nodes of each airport and ACC.

b) Airport aggregation matrix: Airports that connect to the same ACC are merged into an aggregated node. An aggregated-node matrix is built and each node represents a set of airports.

c) Aggregated flight path matrix: Based on the flight path matrix and the aggregated-node matrix, an aggregated flight path matrix is built.

3) TMF estimation: A matlab code ‘linprog’ was used to maximize of traffic flows at 67 busy airports. The default algorithm of the code is interior-point-method which is used to solve linear and nonlinear convex optimization problems [31].

4) Mixed testing: This includes bottom-up and top-down testing. Each sub-model was tested to assure the internal consistency of the model [32].

IV. RESULTS AND DISCUSSION

The results of maximum flow estimation introduced in this section are presented in the order of network, airports and ACCs. In addition to presenting the maximum flows, the comparison between the theoretical prediction of Little’s theorem and ATFM delays is illustrated in the last subsection.

A. Network

The comparison between EMF and TMF at airports, in ACCs and in network is listed in Table II. The ratio of EMF to TMF in airspace is relatively higher than the same ratios at airports and in network. It implies that there is less spare capacity in the en-route airspace than at airports and in network. Hence traffic congestion in airspace will be more severe than at airports if the increase of air traffic demands continues.

<table>
<thead>
<tr>
<th>Traffic</th>
<th>Airports</th>
<th>ACCs</th>
<th>Network</th>
</tr>
</thead>
<tbody>
<tr>
<td>Flights</td>
<td>51.0</td>
<td>121.2</td>
<td>28.8</td>
</tr>
<tr>
<td>EMF/ TMF</td>
<td>42%</td>
<td>56%</td>
<td>42%</td>
</tr>
</tbody>
</table>

The comparison between the EMF and the TMF on all constituent nodes and linear regression analysis are presented in Fig. 3.

The correlation coefficient is 0.96 and the regression coefficients are 0.60 and -142. This relatively high correlation coefficient suggests a very strong dependence between EMF and TMF and the regression analysis suggests that the EMF can be predicted by using the TMF.

In addition to predicting the EMF, the regression coefficients in (9) can be used to quantify the influence of capacity factors. A robust method for quantifying capacity factors will be an important future feature of this research.

\[
\text{EMF}_{\text{Network}} = 0.6 \times \text{TMF}_{\text{Network}} - 142
\]  

B. Airport Operation

1) Comparison between TMF and EMF

All of busy airports are identified as bottlenecks and run out their capacities after maximizing the traffic flows. Figures 4 and 5 illustrate the TMF, EMF and regression analysis at busy and aggregated airports respectively. Despite the relatively
high correlation coefficients of 0.79 and 0.74, there is considerable fluctuation in the TMF curves.

This is an inherent shortcoming of this LP model. The daily capacity is calculated by multiplying hourly capacity and operational hours together. Direct multiplication creates static daily capacities and high level of errors may occur when comparing to the empirical traffic that was operated dynamically. The approach to overcome this shortcoming is introduced in section six.

\[ \text{EMF}_{\text{Busy Airports}} = 0.65 \times \text{TMF}_{\text{Busy Airports}} - 65 \]  
(10)

The regression coefficients in (10) and (11) are expected to be used to quantify capacity factors.

\[ \text{EMF}_{\text{Aggregated Airports}} = 0.25 \times \text{TMF}_{\text{Aggregated Airports}} + 38 \]  
(11)

2) Lagrange Multipliers

The \( \lambda \) at busy airports is nearly to 1 and indicates that the capacity increase at a busy airport results in an equal increase in network capacity. This result has been verified by using a test of increasing capacity at busy airports.

The importance of bottlenecks can be assessed by comparing the \( \lambda \) of the bottleneck-nodes. Fig. 6 depicts the \( \rho \) and \( \lambda \) over the busy airports.

The correlation coefficient is 0.93 and this high value suggests that the priority of investing network capacity should be in line with the level of utilization at the busy airports.

However, the influence of other factors on Lagrange Multipliers needs to be investigated. Using the Lagrange Multiplier as an indicator to assess the importance of bottlenecks requires further research.

\[ \text{EMF}_{\text{ACCs}} = 0.63 \times \text{TMF}_{\text{ACCs}} - 191 \]  
(12)

C. ACC Flight

A very strong dependence between EMF, TMF and ACC capacities is shown in Fig. 7. These high correlation coefficients suggest that the air traffic in Europe is strongly in line with the ACC capacities.

The regression analysis suggests that European ACCs are currently operating at the level of 63\% of TMF. The regression coefficients in (12), 0.63 and -191, can be used to quantify the relative degradation-parameters in future research.

D. ATFM Delays

According to Little’s theorem, the averaged queue length in an M/M/1 queuing system (\( L_s \)) can be formulated as:

\[ L_s = \rho / (1 - \rho) \]  
(13)

In the case of airports and ACCs, the queue lengths equal to ATFM delays multiplied by airport or ACC capacities.

Fig. 8 indicates that the queue lengths at the airports and in the ACCs seem to fit relatively well to the theoretical predicted curve. This suggests that queuing theory has the potential to quantify the influence of capacity factors.
V. FINDINGS

This section provides a summary of the main findings from previous sections.

A. Validation of the LP model

- The relatively high correlation between the EMF and TMF is evidence for the validity of the LP model.
- The high correlation coefficients in Fig. 6 suggest that the LP model can effectively predict the marginal benefit of capacity gain on bottlenecks.

B. Influences of Capacity Factors

We assume that the gap between the EMF and the TMF is caused by the inefficient performance of network capacity factors. The regression coefficients in equations (9) to (12) can potentially be used to quantify the capacity factors.

C. Quantification of Capacity Factors

The curve in Fig. 8 suggests that queuing theory has the potential to translate ATFM delays into capacity. The causes of ATFM delays [29] are recorded by EUROCONTROL and can be quantified by using an inverse model [33]. These causes can also be mapped to the correlated capacity factors.

VI. LIMITATION AND FUTURE WORK

Our LP model is based on the published capacities of airports and ACCs and the traffic demand patterns. Several limitations, their appropriate resolution and several future avenues of research are identified in this section.

A. Inherent Limitations

1) Capacities

In Europe, the ACC capacities are mainly calculated by using the Network Estimation & Visualization of ACC Capacity tool (NEVAC), Reverse CASA and ACCESS. The airport capacities are calculated by using Commonly Agreed Methodology for Airport Capacity Assessment (CAMACA) [20]. The accuracy of these methods has significant impacts on the accuracy of our LP model and as a result a detailed review of these methods is required.

2) Static Nature

This model is set to calculate the daily maximum flow. The results outlined in this paper are valid for the traffic flows in a saturated and static network. The daily capacities are calculated by direct multiplying hourly capacities and operational hours together. The results were validated by comparing them to the empirical traffic data, which were recorded dynamically. Comparing static results to dynamic empirical data cannot reflect the dynamic nature of the European air traffic network and hence there is a need to build a dynamic model.

3) Traffic Demand Pattern

The traffic on 1 July 2012 is used to build a matrix that represents the traffic demand pattern in Europe. Although the demand pattern can be reflected by using congested traffic, the temporal change of demands is not included. More data of traffic profiles on different days are required to enable such generalized demand patterns. In addition to calibrating the demand patterns, more data of traffic profiles can be used to conduct cross validation to our model.

B. Future Research

In order to improve the accuracy of our LP model, the first task of future research is to overcome the inherent limitations of the model by using the proposed resolution.

After improving the LP model, the second task is to map the capacity factors [16, 25] into ATFM delay causes. This enables us to quantify the capacity factors by using queuing theory and regression analysis.

Finally, the contribution of SESAR can be assessed by mapping new operational improvements of SESAR to capacity factors.

VII. CONCLUSION

This paper has developed, for the first time, a linear programming to estimate maximum flows in the European air traffic network. The results suggest that this LP approach is relatively capable to model the air traffic in Europe.

In addition, the influence of the capacity factors is assumed to be linear degradation-parameters and the linear relation between EMF and TMF can be used to quantify these parameters.

Finally, ATFM delays and queuing theory can potentially be used to quantify capacity factors.

ACKNOWLEDGEMENT

The authors thank the R.O.C. Air Force Academy and the Lloyd’s Register Foundation for sponsoring this research. We also thank EUROCONTROL and Mr. Mark Sigerson, Mr. Patrick Tasker and Mr. Nicolas Bruno who generously provided the air traffic data.

REFERENCES
