

## Generazione delle curve di riempimento delle stazioni ferroviarie<sup>(1)</sup>

### Generation of filling curves of railway stations<sup>(1)</sup>

Prof. Ing. Gabriele MALAVASI<sup>(\*)</sup>, Dott. Ing. Carlo Domenico RONZINO<sup>(\*\*)</sup>

**Sommario** - Il metodo denominato “delle curve di riempimento” permette, con un approccio sintetico, cioè non legato alla topologia della stazione ma ad alcuni parametri “globali” quali il numero di binari, la progettazione o la verifica della capacità di sosta delle stazioni ferroviarie, associata ad un prefissato livello di qualità del servizio.

Lo scopo della memoria è descrivere alcuni metodi per la costruzione di curve di riempimento delle stazioni ferroviarie, in funzione, oltre che del livello di qualità atteso, anche del volume di traffico e della legge di distribuzione degli arrivi.

#### 1. Premessa metodologica

Il presente lavoro vuole dare un contributo alle problematiche relative alla valutazione della capacità di sosta delle stazioni ferroviarie, comunque complesse.

Si basa sullo studio analitico della legge di distribuzione degli arrivi e, tenendo conto dei tempi di occupazione dell'infrastruttura da parte delle singole circolazioni, fornisce indicazioni preliminari sulla capacità della stazione.

La sua indipendenza dall'orario lo rende applicabile, fin dalle fasi progettuali preliminari, per ottenere indicazioni sull'opportunità di sviluppare approfondimenti progettuali utilizzando modelli di simulazione più dettagliati che, a fronte di risultati più specifici, richiedono però informazioni più dettagliate (topologia dell'impianto, velocità di percorrenza degli itinerari, livello di automazione dell'apparato di comando e controllo, compilazione dei Prospetti dei binari di arrivo e partenza, ecc.). Per l'uso dei metodi di simulazione si vedano, ad esempio, i

**Summary** - The method of “filling curves” allows the design or verification of the stop capacity of railway stations, associated with a fixed service quality level, with a synthetic approach, i.e. not related to the topology of the station but to some “global” parameters such as the number of tracks.

The purpose of the brief is to describe some methods for constructing filling curves of railway stations as functions of both the expected quality level and the traffic volume and the arrival distribution law.

#### 1. Methodological approach

This work aims at contributing to the analysis of problems concerning the evaluation of stop capacity of railway stations, however complex.

It is based on the analytical study of the arrival distribution law and, taking into account the timing of infrastructure occupation of the individual traffic flows, it provides preliminary indications on the capacity of the station.

Its independence from time makes it applicable, since the preliminary design stages, to obtain guidance on whether to develop design insights using more detailed simulation models that, against more specific results, however require more detailed information (system topology, routes travel speed, automation level of the command and control apparatus, compiling of arrival and departure tracks schedules, etc.). For the use of simulation methods see, for example, the contributions of CORAZZA – ROTA [1], MALAVASI – RICCI [2], CROTTI – DEFLORIO – ELIA – IONCOLI [3] and MALAVASI – MARINI – PETRILLI [4].

(\*) Ordinario di Trasporti presso l'Università di Roma “La Sapienza”.

(\*\*) Dottore di Ricerca in Infrastrutture e Trasporti.

<sup>(1)</sup> Il presente lavoro trae spunto da alcuni risultati della tesi di dottorato di ricerca in Ingegneria delle infrastrutture e dei trasporti, presso l'Università di Roma “La Sapienza” (XXIII ciclo), svolta da C.D. RONZINO.

(\*) Professor of Transportation at the University of Rome “La Sapienza”.

(\*\*) PhD on Infrastructures and Transports.

<sup>(1)</sup> This study is drawn from some results of the doctorate research in infrastructure and transportation engineering at the University of Rome “La Sapienza” (XXIII cycle), carried out by C.D. RONZINO.

contributi di CORAZZA – ROTA [1], MALAVASI – RICCI [2], CROTTI – DEFLORIO – ELIA – IONCOLI [3] e MALAVASI – MARINI – PETRILLI [4].

Il metodo oggetto dello studio fornisce il numero minimo  $n$  di binari di stazionamento che debbono comporre il piano di stazione per soddisfare, con un predefinito livello di qualità complessiva del servizio, gli arrivi da una o più direzioni, nell'ipotesi che essi possano essere utilizzati indipendentemente uno dall'altro, senza cioè vincoli impiantistici al loro utilizzo.

Tale ipotesi si ritiene non costituisca una limitazione, in quanto è sempre possibile suddividere l'impianto di stazione in più parti, per ciascuna delle quali valga la predetta ipotesi.

### 1.1. Richiami di teoria delle curve di riempimento

L'utilità delle curve di riempimento nell'ingegneria ferroviaria, in particolare per la valutazione della capacità di sosta delle stazioni, è stata anche recentemente dimostrata da GIOVINE - LEONARDI [5] e DE VITA - GIOVINE - RICCI [6], con applicazione a casi concreti.

Con la presente memoria si vuole dare un ulteriore contributo, da un lato, dimostrando la possibilità, in certe condizioni, di scrivere l'equazione risolutiva in forma chiusa di insiemi (famiglie) di curve di riempimento, dall'altro, proponendo una metodo statistico per elaborarle rendendo minimo l'errore commesso.

Le curve di riempimento consentono una rappresentazione sintetica delle funzioni di sosta di una stazione ferroviaria, quando siano noti la distribuzione degli arrivi dei treni nell'arco di un periodo di riferimento e il loro tempo di sosta medio.

Il problema che si pone ed al quale si dà soluzione è il seguente: *nota la distribuzione degli arrivi di  $N$  treni in una stazione e il loro tempo di sosta, quanti binari devono essere previsti nella stazione affinché gli  $N$  treni possano svolgere il loro servizio garantendo un prefissato livello di qualità?*

Tale problema si può presentare in una duplice veste, come "progetto" del numero dei binari di una nuova stazione o come "verifica".

Nel primo caso la stazione non esiste e quindi le distribuzioni degli arrivi dei treni e dei loro tempi di sosta, non essendo osservabili, possono essere conosciute solo per similitudine oppure imposte come dati di progetto; nel secondo caso invece, poiché la stazione esiste, si ha una conoscenza diretta della distribuzione degli arrivi dei treni e dei loro tempi di sosta, che può essere deterministica o probabilistica (qualora si possa costruire una legge di distribuzione dei ritardi o degli anticipi dei treni rispetto all'orario programmato).

Si consideri il fenomeno aleatorio costituito da una successione di  $N$  treni, nel tempo di riferimento  $T$ .

All'interno di tale insieme si considerino  $n$  arrivi; la varia-

*The method object of the study provides the minimum number  $n$  of stabling tracks that must compose the station plan to meet arrivals from one or more directions with a pre-set level of overall service quality, assuming that they could be used independently from each other, without system constraints on their use.*

*This hypothesis is deemed not to constitute a restriction, since the station plant can always be split into several parts, for each of which the aforementioned hypothesis is valid.*

#### 1.1.1. References of theory of filling curves

*The utility of filling curves in railway engineering, in particular for the assessment of the stop capacity of stations, was also recently demonstrated by GIOVINE - LEONARDI [5] and DE VITA - GIOVINE - RICCI [6], with application to concrete cases.*

*With this work we want to give a further contribution on the one hand, demonstrating the possibility, under certain conditions, to write the equation solution in closed form (families) of sets of filling curves, on the other hand, proposing a statistical method to process them reducing the error made to a minimum.*

*Filling curves provide a concise representation of the stop functions of a railway station, when the train arrivals distribution over a reference period and their average stop time is known.*

*The problem that arises, and to which a solution is given is as follows: the distribution of the arrivals of  $N$  trains at a station and their stop time being known, how many tracks must be provided in the station for the  $N$  trains in order to perform their service ensuring a pre-set quality level?*

*This problem may occur in a dual role, as "project" of the number of tracks of a new station or as "verification".*

*In the first case, there is no station and therefore the train arrivals distributions and their stop times, that are unobservable, can only be known for similarity or can be imposed as project data; in the second case instead, since the station exists, there is direct knowledge of the distribution of train arrivals and their stop times, that can be deterministic or probabilistic (where a delay or advance distribution law of trains compared to programmed time can be developed).*

*Just consider the random phenomenon consisting of a sequence of  $N$  trains, in the  $T$  reference time.*

*Within that set  $n$  arrivals are considered; the random variable describing the phenomenon is  $\tau_n$ , the time interval during which  $n$  arrivals occur. The specific  $t_n$  values that the random variable  $\tau_n$  assumes are defined by a  $P_n$  continuous probability distribution function (fig. 1).*

$$P_n(t_n) = P_n(\tau_n \leq t_n) = \int_0^{t_n} p(\tau, n) d\tau \quad (1)$$

*The  $P_n(t_n)$  value represents, by definition, the probability of  $n$  arrivals occurring in a time not greater than  $t_n$ , being  $p(\tau, n)$  the probability density of  $n$  arrivals according to  $\tau$ .*

bile aleatoria che descrive il fenomeno sia  $\tau_n$ , l'intervallo di tempo durante il quale avvengono gli  $n$  arrivi. Gli specifici valori  $t_n$  che la variabile aleatoria  $\tau_n$  assume siano definiti da una funzione continua di distribuzione di probabilità  $P_n$  (fig. 1).

$$P_n(t_n) = P_n(\tau_n \leq t_n) = \int_0^{t_n} p(\tau, n) d\tau \quad (1)$$

Il valore  $P_n(t_n)$  rappresenta per definizione la probabilità che gli  $n$  arrivi avvengano in un tempo non maggiore di  $t_n$ , essendo  $p(\tau, n)$  la densità di probabilità di  $n$  arrivi in funzione di  $\tau$ .

Detto valore, qualora  $t_n$  corrisponda al tempo di servizio per treno associato a ciascun binario della stazione, rappresenta una misura della probabilità di funzionamento irregolare della stazione; infatti se gli  $n$  arrivi si succedessero in un tempo inferiore o, al più, uguale al valore del tempo di servizio, nella stazione si avrebbe una congestione (fig. 2).

Poniamo dunque:

$$P_n(t_n) = \sigma \quad (2)$$

La probabilità che gli  $n$  arrivi avvengano in un tempo maggiore di  $t_n$ , cioè, per quanto detto, la probabilità che nella stazione non si produca congestione, è (fig. 3):

$$P_n^*(\tau_n \geq t_n) = \int_{t_n}^{\infty} p(\tau, n) d\tau = 1 - \sigma \quad (3)$$

Estendendo queste considerazioni ad ogni  $n$  appartenente alla successione dei numeri naturali si arriva a definire un insieme di punti nel piano di coordinate cartesiane  $n, t$ , ciascuno dei quali, fissato il valore del parametro  $\sigma$ , rappresenta:

- per un determinato  $n$ , definito con probabilità  $(1-\sigma)$ , l'intervallo di tempo minimo  $t$ , durante il quale si hanno gli  $n$  arrivi;
- per un determinato  $t$ , definito con probabilità  $(1-\sigma)$ , il numero massimo di arrivi.

L'insieme di punti così determinato è denominato "curva di riempimento" (fig. 4); la sua rappresentazione grafica convenzionale è una spezzata o una curva continua (fig. 5), entrambe comunque non decrescenti, poiché il fenomeno fisico che si descrive è una successione ordinata di arrivi.

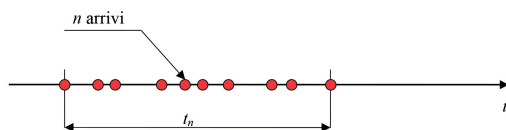


Fig. 1 - Distribuzione di  $n$  arrivi nel tempo  $t_n$ .  
Fig. 1 - Distribution of  $n$  arrivals over  $t_n$  time.

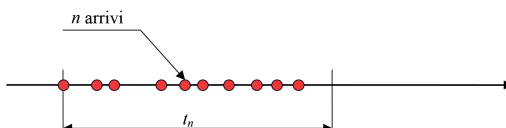


Fig. 2 - Distribuzione di  $n$  arrivi in un tempo inferiore a  $t_n$ ; situazione di probabile congestione.  
Fig. 2 - Distribution of  $n$  arrivals in a time lapse less than  $t_n$ ; probable congestion situation.

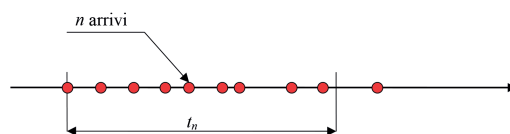


Fig. 3 - Distribuzione di  $n$  arrivi in un tempo maggiore di  $t_n$ .  
Fig. 3 - Distribution of  $n$  arrivals in a time lapse greater than  $t_n$ .

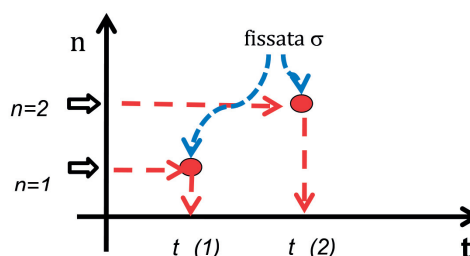


Fig. 4 - Costruzione per punti della curva di riempimento.  
Fig. 4 - Building by points of the filling curve ( $\sigma$  fixed).

This value, if  $t_n$  is the service time per train associated with each station's track, represents a measure of the probability of irregular operation of the station; in fact if the  $n$  arrivals were to succeed in a time lapse lower than or, at the most, equal to the value of the service time, there would be a congestion in the station (fig. 2).

Let us therefore suppose:

$$P_n(t_n) = \sigma \quad (2)$$

The probability that  $n$  arrivals occur in a time lapse greater than  $t_n$  i.e., as has been said, the probability that the station does not produce congestion, is (fig. 3).

$$P_n^*(\tau_n \geq t_n) = \int_{t_n}^{\infty} p(\tau, n) d\tau = 1 - \sigma \quad (3)$$

By extending these considerations to every  $n$  belonging to the sequence of natural numbers the definition of a set of points in the Cartesian coordinate plane  $n, t$ , is obtained, each of which, having set the value of parameter  $\sigma$ , represents:

- for a given  $n$ , defined with probability  $(1-\sigma)$ , the minimum time interval  $t$ , during which we have the  $n$  arrivals;
- for a given  $t$ , defined with probability  $(1-\sigma)$ , the maximum number of arrivals.

The set of points so determined is referred to as "filling curve" (fig. 4); its conventional graphic representation is a dashed or a continuous curve (fig. 5), both however not decreasing, since the physical phenomenon that it describes is an ordered sequence of arrivals.

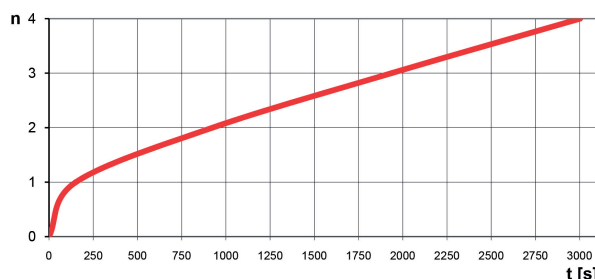


Fig. 5 - Esempio di curva di riempimento (con rappresentazione di curva continua).

Fig. 5 - Filling curve example (with continuous curve representation).

Una curva di riempimento così costruita si può considerare come una successione di arrivi ordinata per intervalli via via crescenti, soggetti alla probabilità  $\sigma$  che possa verificarsi una successione più sfavorevole, cioè con intervalli più ravvicinati.

Al variare del parametro  $\sigma$  si ottiene un insieme di curve di riempimento (famiglia di curve di riempimento), ciascuna delle quali esprime il funzionamento della stazione in relazione al dato valore di qualità del servizio.

### 1.2. Applicazione delle curve di riempimento nelle fasi di progetto e di verifica della capacità di sosta di una stazione ferroviaria

L'applicazione delle curve di riempimento nella fase di progetto di una nuova stazione o di verifica di una esistente, nell'ipotesi di modifica o incremento delle circolazioni, presuppone la costruzione di una seconda curva, in prima ipotesi parallela alla prima e traslata da essa di un valore  $t_s$ , lungo l'asse dei tempi, che rappresenta il tempo di sosta medio dei treni nella stazione stessa. Il tempo di sosta medio, uguale per ipotesi per ogni  $n$ , dipende dal tempo impiegato dal treno a percorrere l'itinerario di arrivo, dal tempo di servizio espletato sul binario di stazionamento e dal tempo necessario a rendere nuovamente utilizzabile il binario di stazionamento dopo la ripartenza del treno. Inoltre nel computo generale si deve introdurre anche un termine che stima il ritardo medio causato dall'interdizione dei movimenti incompatibili - di arrivo e di partenza - con quello del treno in esame [8, 9, 10, 11, 12, 13].

Tale seconda curva è chiamata *curva di svuotamento*, poiché rappresenta la successione ordinata delle partenze di treni dalla stazione.

Per la fase di progetto, una coppia di curve di riempimento e di svuotamento (fig. 6) è utilizzata per:

- ricavare il numero di binari necessari che garantisce il rispetto del livello di qualità voluto, noto il tempo di sosta medio;
- ricavare il tempo di sosta medio che consente il servizio di  $n$  treni nel rispetto del livello di qualità voluto, noto il numero di binari.

A filling curve built this way can be considered as a sequence of arrivals sorted by gradually increasing intervals, subject to the probability  $\sigma$  that a more unfavourable succession can occur, with closer timings.

As the  $\sigma$  parameter changes a filling curves set is achieved (filling curves family), each of which expresses the operation of the station in relation to the given service quality value.

### 1.2. Application of filling curves in the project and testing phases of the stop capacity of a railway station

The application of filling curves in the design phase of a new station or verification of an existing one, in the case of changes or increases of traffic flows, involves the construction of a second curve, at first parallel to the first curve and shifted from it by a  $t_s$  value, along the axis of time, representing the average stop time of the trains at the station itself. The average stop time, assumed equal for every  $n$ , depends on the time the train takes to run the arrival route, on the service time carried out on the stop track and on the time necessary to make the track usable again after the departure of the train. In addition, in the overall calculation a term that estimates the average delay caused by the locking of incompatible movements - of arrival and departure - should also be introduced with that of the train concerned [8, 9, 10, 11, 12, 13].

This second curve is called the emptying curve, because it represents the orderly succession of departures of trains from the station.

For the design phase, a pair of filling and emptying curves (fig. 6) is used to:

- obtain the number of tracks required that ensures compliance with the desired quality level, being the average stop time known;
- obtain the average stabling time allowing the service of  $n$  trains in compliance with the desired quality level, being the number of tracks known.

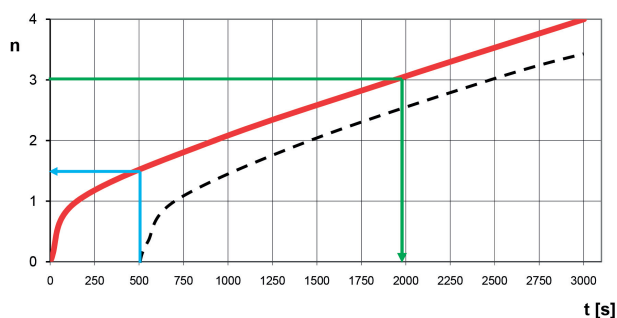


Fig. 6 - Esempio di curva di riempimento e, a tratteggio, di curva di svuotamento.

Fig. 6 - Example of filling curve and, dashed, of emptying curve.

Analogamente per la fase di verifica le suddette curve sono utilizzate, noti il tempo di sosta medio e il numero di binari, per verificare che il tempo di sosta reale sia congruente con il valore di qualità voluto o per verificare che il numero di binari esistente sia congruente con il valore di qualità voluto.

I metodi di progetto e verifica possono anche essere utilizzati per progettare o verificare, fissati gli altri valori, la distribuzione degli arrivi, cioè il programma di esercizio assegnato alla stazione, sia reale (stazione in esercizio), sia teorico (stazione non in esercizio o in esercizio ma con un altro programma), oppure per supportare un'analisi costi-benefici.

Ciò può essere utile, ad esempio, quando è necessario considerare nell'offerta ferroviaria il livello di qualità atteso, in termini assoluti (come probabilità di arrivi compresi in un certo intervallo di tempo) o in termini relativi, rispetto ad un livello di riferimento standard.

## 2. Metodologie per la costruzione delle curve di riempimento

In generale la costruzione di una curva di riempimento o, al variare del parametro  $\sigma$ , di una famiglia di curve di riempimento può avvenire in tre modi:

- se si dispone di una successione di arrivi di treni: *procedimento numerico diretto*;
- se non si dispone di una successione di arrivi ma:
  - si conosce la funzione di probabilità degli arrivi di singoli treni: *procedimento teorico-statistico*;
  - si conosce la funzione di probabilità degli arrivi di enuple di treni: *procedimento teorico-sintetico*.

Il procedimento teorico-sintetico si basa quindi sulla conoscenza della funzione densità di probabilità  $p(\tau, n)$  della (1).

### 2.1. Procedimento numerico diretto

Con questo procedimento si determina, per ogni  $n$ , il tempo  $t_n$  in funzione di  $\sigma$  (che indicheremo con  $t_n(\sigma)$ ), a partire da una successione nota degli arrivi dei treni, cioè dall'orario ferroviario. Per arrivo di un treno si intende l'istante in cui il treno riceve l'ordine di ingresso in stazione; ciò equivale a dire che l'insieme degli arrivi viene ordinato in  $m$  classi temporali sulla base degli intertempi fra arrivi di  $n$  treni.

Stabilita l'ampiezza delle classi temporali l'insieme degli arrivi consente la costruzione, per ciascuno dei valori assegnati a  $n$ , dei seguenti vettori:

- 1) vettore  $I_n$  degli intertempi, cioè degli intervalli fra gli arrivi di  $n$  treni (come vedremo più avanti la sua determinazione può avvenire secondo un procedimento con "ripetizione" o "non ripetizione" delle estrazioni);
- 2) vettore  $R_n$  della frequenza relativa degli intertempi, il

*Similarly for the verification phase, these curves are used, with the average stop time and the number of tracks known, to verify that the actual stop time is congruent with the desired quality value or to verify that the existing number of tracks is congruent with the desired quality value.*

*The design and verification methods may also be used, having set other values, to design or test the distribution of arrivals, i.e. the operation programme assigned to the station, both real (station in operation) and theoretical (non-operating station or in operation but with another programme), or to support a cost-benefit analysis.*

*This can be useful, for example, when it is necessary to consider the expected quality level in the railway offer in absolute terms (as probability of arrivals within a certain time interval) or in relative terms compared to a standard reference level.*

## 2. Methodologies for the construction of filling curves

*In general, the construction of a filling curve or, as the parameter varies, of a family of filling curves can be performed in three ways:*

- *if there is a sequence of arrivals of trains: direct numerical procedure;*
- *if there is no sequence of arrivals but:*
  - *the probability function of the arrivals of individual trains is known: theoretical-statistical procedure;*
  - *if the probability function of arrivals of tuple trains is known: theoretical-synthetic procedure.*

*The theoretical-synthetic procedure is therefore based on the knowledge of the probability density function  $p(\tau, n)$  of (1).*

### 2.1. Direct numerical procedure

*With this procedure, for each  $n$ , the  $t_n$  time as a function of  $\sigma$  (which we will indicate with  $t_n(\sigma)$ ), is determined starting from a known sequence of train arrivals, i.e. from the railway timetable. The meaning of train arrival is the moment when the train receives the order to enter the station; this is equivalent to saying that the arrivals set is ordered in  $m$  temporal classes based on the split times between arrivals of  $n$  trains.*

*Having set the range of temporal classes, the set of arrivals allows the construction of the following vectors, for each of the values assigned to  $n$ :*

- 1) *vector  $I_n$  of split times, i.e. of intervals between arrivals of  $n$  trains (as we will see later on its determination can occur according to a process with "repetition" or "non-repetition" of extractions);*
- 2) *vector  $R_n$  of the relative frequency of split times, whose generic element  $b_i$  expresses the portion of the total, for the generic temporal class;*

cui elemento generico  $b_i$  esprime, per la generica classe temporale, la quota parte sul totale;

- 3) vettore  $C_n$  della frequenza cumulata degli intertempi, costruito a partire dal vettore  $R_n$ , il cui elemento generico  $a_i = \sum b_i$ .

La conoscenza dei tre vettori consente, fissato il valore del parametro  $\sigma$ , di calcolare, per ogni  $n$ , il valore del tempo  $t_n(\sigma)$  relativo:

$$t_n(\sigma) = t_{i-1} + \frac{(\sigma - a_{i-1})(t_i - t_{i-1})}{a_i - a_{i-1}}$$

dove  $t_{i-1}$  è il valore, espresso in minuti, della classe temporale cui è associato un valore di frequenza cumulata  $a_{i-1}$  che approssima per difetto il valore del parametro  $\sigma$ , e  $(t_i - t_{i-1})$  è l'ampiezza della classe temporale.

Ciascuna coppia di valori  $(n; t_n(\sigma))$  individua sul piano di coordinate cartesiane  $n, t$  un punto appartenente alla curva di riempimento.

Nel calcolo degli intertempi fra arrivi di  $n$  treni è possibile operare in due modi diversi, in analogia al campionamento con "ripetizione" delle estrazioni o "senza ripetizione".

Nel primo caso, ogni ennupla successiva ad una data si forma a partire dall'elemento  $(i+1)$ -esimo, dove  $i$  è il primo elemento dell'ennupla data; nel caso invece "senza ripetizione", il primo treno dell'ennupla successiva a quella data è l'elemento  $(i+n+1)$ -esimo; in altre parole:

- nel caso con ripetizione ogni ennupla di treni in arrivo ha  $(n-1)$  elementi (cioè treni) in comune con l'ennupla precedente;
- nel caso "senza ripetizione" le ennuple di treni in arrivo non hanno alcun elemento in comune fra loro.

Dal punto di vista del funzionamento della stazione appare più corretto operare secondo il modello "con ripetizione"; infatti il sistema di esercizio è tale che un binario, appena si libera, può essere occupato da un treno in attesa al segnale di protezione, indipendentemente dal fatto che questo treno in attesa appartenga ad una ennupla successiva.

La costruzione senza ripetizione delle estrazioni dei valori fa sì che le ennuple vengano considerate singolarmente; ciò equivale a dire che il primo treno dell'ennupla successiva può entrare in stazione solo dopo che l'ultimo della precedente ennupla ha liberato il binario di stazionamento. Ad esempio, per una successione di sei treni, designati con le lettere A, B, C, D, E e F, nel caso di estrazione di coppie con ripetizione si avrebbe: AB, BC, CD, DE, EF; nel caso di estrazione senza ripetizione invece solo: AB, CD, EF.

## 2.2. Procedimento teorico-statistico

Il procedimento teorico-statistico rappresenta un'applicazione del procedimento numerico diretto all'insieme di arrivi definito da un modello di arrivi.

- 3) vettore  $C_n$  of the cumulated frequency of split times, built starting from vector  $R_n$ , whose generic element  $a_i = \sum b_i$ .

Having set the value of the parameter, the knowledge of the three vectors allows calculating the value of the relative time  $t_n(\sigma)$ , for each  $n$ ,

$$t_n(\sigma) = t_{i-1} + \frac{(\sigma - a_{i-1})(t_i - t_{i-1})}{a_i - a_{i-1}}$$

where  $t_{i-1}$  is the value, in minutes, of the temporal class that is associated with a cumulative frequency value  $a_{i-1}$  which under approximates the  $\sigma$  parameter value, and  $(t_i - t_{i-1})$  is the range of the temporal class.

Each pair of values  $(n; t_n(\sigma))$  identifies a point in the filling curve on the Cartesian coordinate plane  $n, t$ .

In calculating the split times between arrivals of  $n$  trains it is possible to operate in two different ways, in analogy to sampling with "repetition" of the extractions or "without repetition".

In the first case, each tuple subsequent to a given one is formed starting from the element  $(i+1)$ -th, where  $i$  is the first element of the tuple given; in case of "no repetition", the first train of the next tuple to the one given is the element  $(i+n+1)$ -th; in other words:

- if there is repetition, every tuple of incoming trains has  $(n-1)$  elements (i.e. trains) in common with the previous tuple;
- if there is "no repetition" the tuples of incoming trains have nothing in common with each other.

From the point of view of the operation of the station it appears to be more appropriate to operate according to the "repetition" model; in fact, the operating system is such that as soon as a track is clear, it can be occupied by a train waiting at the entry signal, regardless of whether this waiting train belongs to a subsequent tuple.

Construction of the values without repetition of extractions causes the tuples to be considered individually; this means that the first train of the subsequent tuple can enter the station only after the last of the previous tuple has cleared the stop track. For example, for a succession of six trains, named by the letters A, B, C, D, E and F, in the case of extraction of repeating pairs we would have: AB, BC, CD, DE, EF; in the case of extraction without repetition instead we would only have: AB, CD, EF.

## 2.2. Theoretical-statistical procedure

The theoretical-statistical procedure is an application of the direct numerical procedure to a set of arrivals defined by an arrival model.

In this study we used the Poisson arrivals model, expressed by a law of exponential distribution; this model, typical of transport engineering and confirmed by the analysis of the split times referred to in paragraph 3.2, is

Nel presente studio è stato utilizzato il modello di arrivi di *Poisson*, espresso da una legge di distribuzione di tipo esponenziale; questo modello, tipico dell'ingegneria dei trasporti e confermato dall'analisi degli intertempi di cui al paragrafo 3.2, si basa su un solo parametro statistico, il tasso medio  $\lambda$  di arrivi.

In funzione di  $\lambda$  si è provveduto a "generare" casualmente, mediante la funzione *random*, le diverse successioni di arrivi e la curva cumulata della distribuzione di probabilità.

Ognuno dei numeri reali generati dalla funzione *random* rappresenta il valore della distribuzione di probabilità esponenziale:

$$y_i = 1 - e^{-\lambda t_i}$$

da cui si ricava dunque il valore della variabile  $t_i$  "istante di arrivo dell' $i$ -esimo treno":

$$t_i = \frac{-\ln(1 - y_i)}{\lambda}$$

L'insieme dei tempi  $t_i$  costituisce quanto cercato, cioè la successione degli arrivi nel tempo di riferimento (orario ferroviario).

Rimane da stabilire il numero minimo di iterazioni necessario per avere un risultato statisticamente significativo; ciò costituisce un tipico problema di stima campionaria della media.

Il numero minimo si ricava una volta stabilito l'errore massimo ammesso nella stima della media, funzione dell' $m$ -esima iterazione (cioè dell' $m$ -esimo orario generato casualmente).

Detta quindi  $\mu$  la media della popolazione,  $\bar{X}$  la media stimata del campione di  $m$  elementi,  $s_m$  la deviazione standard del campione di  $m$  elementi<sup>(2)</sup> e  $t_\alpha$  il coefficiente della distribuzione di Student relativo all'intervallo di confidenza  $\alpha$ , si può scrivere:

$$\mu = \bar{X} \pm t_\alpha \frac{s_m}{\sqrt{m}} \quad (4)$$

La (4) rappresenta una stima della media con un campione di  $m$  elementi (cioè  $m$  curve di riempimento generate), quando non sia nota la varianza della popolazione [7]. Il termine  $t_\alpha \frac{s_m}{\sqrt{m}}$  misura l'errore che si commette all' $m$ -esima iterazione.

Detto  $d$  l'errore massimo ammesso, stabilito a priori, per verificare se il numero  $m$  di iterazioni è sufficiente, cioè se l'errore commesso nella stima della media, con un

based on a single statistical parameter, the average rate of arrivals  $\lambda$ .

According to  $\lambda$  we proceeded to randomly "generate" the different successions of arrivals and the cumulative curve of probability distribution, using the random function.

Each of the real numbers generated by the random function represents the exponential probability distribution value:

$$y_i = 1 - e^{-\lambda t_i}$$

from which the value of the variable  $t_i$  "arrival time of the  $i$ -th train" is derived":

$$t_i = \frac{-\ln(1 - y_i)}{\lambda}$$

All the  $t_i$  times constitute what is researched, i.e. the succession of arrivals in the reference time (train timetable).

What must still be determined is the minimum number of iterations needed to have a statistically significant result; this is a typical problem of sample estimate of the average.

The minimum number is obtained once the maximum permissible error in the estimate of the average is known, according to the  $m$ -th iteration (i.e. the  $m$ -th time randomly generated).

Therefore given  $\mu$  as the average population,  $\bar{X}$  the estimated average of  $m$  elements sample,  $s_m$  the standard deviation of  $m$  elements sample<sup>(2)</sup> and  $t_\alpha$  the Student distribution coefficient relating to the confidence interval  $\alpha$ , we can write:

$$\mu = \bar{X} \pm t_\alpha \frac{s_m}{\sqrt{m}} \quad (4)$$

(4) represents an estimate of the average with a sample of  $m$  elements (i.e.  $m$  filling curves generated), when the population variance is not known [7]. The term  $t_\alpha \frac{s_m}{\sqrt{m}}$  measures the error that is made at the  $m$ -th iteration.

Given  $d$  as the maximum allowable error, established a priori, to verify whether number  $m$  of iterations is sufficient, i.e. if the error in estimating the average, with a sample of  $m$  elements, is less than the maximum accepted, one must check whether at the  $m$ -th iteration, the following inequality is satisfied<sup>(3)</sup>:

$$d^2 \geq \frac{s_m^2 t_\alpha^2}{m} \quad (5)$$

<sup>(2)</sup> The variance of the sample of  $m$  elements, or its correct estimate, is:  $s_m^2 = \frac{\sum (X_i - \bar{X})^2}{(m-1)}$

<sup>(3)</sup> The square of the terms expressing errors is a plain trick that simplifies calculations, because it allows working with the term  $m$  and with the sample variance.

<sup>(2)</sup> La varianza del campione di  $m$  elementi, o meglio la sua stima corretta, è:  $s_m^2 = \frac{\sum (X_i - \bar{X})^2}{(m-1)}$

campione di  $m$  elementi, è minore di quello massimo accettato, bisogna verificare se, all' $m$ -esima iterazione, risulta soddisfatta la seguente disuguaglianza<sup>(3)</sup>:

$$d^2 \geq \frac{S_m^2 t_a^2}{m} \quad (5)$$

La (5) esprime il fatto che il quadrato dell'errore commesso deve essere minore del quadrato dell'errore accettato a priori; nel caso in cui la (5) non sia soddisfatta, il numero di iterazioni non è sufficiente.

Il metodo di calcolo è dunque iterativo; di esso può essere data anche un'utile rappresentazione grafica, riportata in fig. 7, nella quale è rappresentato il valore di  $t_n$  ( $n=2$ ) associato alle curve di riempimento che hanno  $\lambda=2$  treni/ora; in essa la linea centrale (nera) rappresenta la stima della media di  $t_n$ , al variare dell'iterazione, la coppia di linee più esterne rappresenta l'intervallo di errore commesso all' $m$ -esima iterazione (linee rosse) e la coppia di linee parallele a quella centrale rappresenta l'errore ammesso (linee verdi).

Come si può vedere, l'errore commesso diventa accettabile (cioè interno all'intervallo  $\pm d$ ) solo alla ventesima iterazione.

La convergenza del metodo è assicurata dal fatto che per  $m \rightarrow \infty$  il secondo membro della (5) tende a 0.

### 2.3. Il procedimento teorico sintetico

Il procedimento teorico sintetico consente il calcolo delle coordinate di punti della curva di riempimento.

Si basa sulla risoluzione della funzione integrale (1) risolvibile analiticamente qualora sia nota la funzione di densità di probabilità  $p(\tau, n)$ .

È questo il caso di arrivi poissoniani, per i quali la funzione cercata è la funzione di Erlang.

La (1) diventa pertanto:

$$\sigma = \int_0^{t_n} \frac{\lambda^n}{(n-1)!} \tau^{n-1} e^{-\lambda\tau} d\tau \quad (6)$$

con  $\lambda > 0$ .

Nella tabella 1 si riassumono alcune soluzioni della (6).

<sup>(3)</sup> L'elevazione al quadrato dei termini esprimenti gli errori è un semplice artificio che semplifica i calcoli, perché permette di lavorare con il termine  $m$  e con la varianza campionaria.

(5) esprime il fatto che la square of the error must be less than the square of the error accepted a priori; in the event that (5) is not met, the number of iterations is not enough.

Hence, the calculation method is iterative; a useful graphical representation of it can also be given, shown in fig. 7, in which the value of  $t_n$  ( $n=2$ ) is represented associated with filling curves that have  $\lambda=2$  trains per hour; in it the centre line (black) represents an estimate of the average of  $t_n$ , as the iteration varies, the outer lines pair represents the range of the error made to the  $m$ -th iteration (red lines) and the pair of parallel lines to the middle one represents the error allowed (green lines).

As you can see, the error becomes acceptable (i.e. within the range  $\pm d$ ) only at the twentieth iteration.

Convergence of the method is ensured by the fact that for  $m \rightarrow \infty$  the second member of (5) tends to 0.

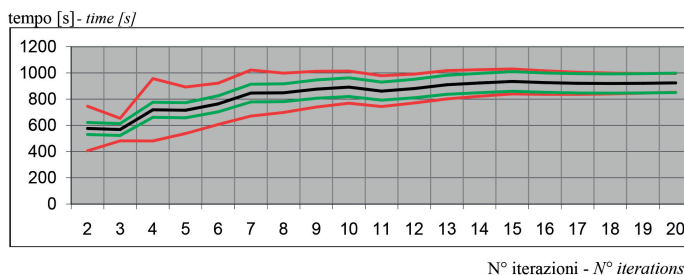


Fig. 7 - Esempio di rappresentazione grafica del calcolo iterativo del valore di  $t_n$ : in ascisse è rappresentato il numero di iterazioni; in ordinate l'errore stimato del valore di  $t_n$  (linee rosse esterne), la media (linea centrale nera) e l'errore ammesso di  $t_n$  (linee verdi interne).

Fig. 7 - Example of a graphical representation of the iterative calculation of the value of  $t_n$ : the x-axis represents the number of iterations; in the ordinate the estimated error of the value  $t_n$  (red external lines), the average (centre black line) and the error allowed for  $t_n$  (internal green lines).

### 2.3. Theoretical-synthetic procedure

The theoretical-synthetic procedure allows the calculation of coordinates of points of the filling curve.

It is based on the resolution of the integral function (1) that is solvable analytically if the probability density function  $p(\tau, n)$  is known.

This is the case for Poisson arrivals, for which the function searched is the Erlang function.

(1) therefore becomes:

$$\sigma = \int_0^{t_n} \frac{\lambda^n}{(n-1)!} \tau^{n-1} e^{-\lambda\tau} d\tau \quad (6)$$

with  $\lambda > 0$ .

Table 1 summarises some solutions of (6).

$t_n$  values are derived, numerically or graphically, from equations that represent solutions of integral functions.

I valori  $t_n$  si ricavano, per via grafica o numerica, dalle equazioni che rappresentano le soluzioni delle funzioni integrali.

### 3. Esempificazione dei procedimenti esposti

In questa sezione sono riportati alcuni esempi numerici dell'applicazione dei tre procedimenti esposti.

#### 3.1. Procedimento numerico diretto

La costruzione delle curve di riempimento è stata svolta con riferimento a stazioni caratterizzate da flussi di traffico diversi (tabella 2).

I tre casi prescelti permettono di coprire, dal punto di vista dei volumi di traffico, gran parte delle situazioni reali.

Le curve di riempimento sono state ricavate fissando tre valori crescenti del parametro  $\sigma$ , ciascuno dei quali misura una *performance*, via via decrescente, della qualità del servizio ferroviario reso:

- $\sigma = 0,03$ ; cioè probabilità del 97% che gli  $n$  arrivi avvengano in un intervallo di tempo maggiore di quello ricavato dal calcolo,  $t_n(\sigma)$ ;
- $\sigma = 0,05$ ; cioè probabilità del 95% che gli  $n$  arrivi avvengano in un intervallo di tempo maggiore di quello ricavato dal calcolo,  $t_n(\sigma)$ ;
- $\sigma = 0,10$ ; cioè probabilità del 90% che gli  $n$  arrivi avvengano in un intervallo di tempo maggiore di quello ricavato dal calcolo,  $t_n(\sigma)$ .

Per semplicità sono stati considerati i valori di  $n$  fino a 4.

Di seguito si riportano i grafici (figg. 8, 9 e 10) e le tabelle 3, 4 e 5, relativi al solo modello di conteggio con ripetizione.

TABELLA 1 – TABLE 1  
SOLUZIONI DELLA FUNZIONE INTEGRALE NEL CASO DELLA FUNZIONE DI ERLANG  
SOLUTIONS OF THE INTEGRAL FUNCTION IN THE CASE OF THE ERLANG FUNCTION

	Funzione integrale <i>Integral function</i>	Soluzione che permette il calcolo di $t_n$ <i>Solution allowing calculation of <math>t_n</math></i>
$n = 1$	$\sigma = \int_0^{t_1} \lambda e^{-\lambda \tau} d\tau$	$t_1 = \frac{\ln(1 - \sigma)}{-\lambda}$
$n = 2$	$\sigma = \int_0^{t_2} \lambda^2 \tau e^{-\lambda \tau} d\tau$	$e^{-\lambda t_2} = \frac{(1 - \sigma)}{1 + t_2 \lambda}$
$n = 3$	$\sigma = \int_0^{t_3} \frac{\lambda^3 \tau^2 e^{-\lambda \tau}}{2} d\tau$	$e^{-\lambda t_3} = \frac{2(1 - \sigma)}{\lambda^2 t_3^2 + 2\lambda t_3 + 2}$
$n = 4$	$\sigma = \int_0^{t_4} \frac{\lambda^4 \tau^3 e^{-\lambda \tau}}{6} d\tau$	$e^{-\lambda t_4} = \frac{6(1 - \sigma)}{\lambda^3} \frac{1}{t_4^3 + \frac{3t_4^2}{\lambda} + 6 \frac{t_4}{\lambda^2} + 6\lambda^{-3}}$
$n$ generico <i>n generic</i>	$\sigma = \int_0^{t_n} \frac{\lambda^n}{(n-1)!} \tau^{n-1} e^{-\lambda \tau} d\tau$	$\frac{(n-1)!}{\lambda^{n-1}} \sigma = -t_n^{n-1} e^{-\lambda t_n} + (n-1) \int_0^{t_n} \tau^{n-2} e^{-\lambda \tau} d\tau$

### 3. Exemplification of proceedings shown

This section provides some numerical examples of the application of the three procedures shown.

#### 3.1. Direct numerical procedure

The construction of filling curves was carried out with reference to stations with different traffic flows (table 2).

The three cases chosen cover most real situations, from the point of view of traffic volumes.

Filling curves were obtained by fixing three increasing values of parameter  $\sigma$ , each of which measures a gradually decreasing performance of the railway service quality rendered:

- $\sigma = 0.03$ ; i.e. 97% probability that  $n$  arrivals occur in a time interval greater than that derived from the calculation,  $t_n(\sigma)$ ;
- $\sigma = 0.05$ ; i.e. 95% probability that  $n$  arrivals occur in a time interval greater than that derived from the calculation,  $t_n(\sigma)$ ;
- $\sigma = 0.10$ ; i.e. 90% probability that  $n$  arrivals occur in a time interval greater than that derived from the calculation,  $t_n(\sigma)$ ;

For simplicity the values of  $n$  up to 4 were considered.

Please find below the charts (figures 8, 9 and 10) and tables 3, 4 and 5, relating only to the counting model with repetition.

TABELLA 2 – TABLE 2  
PROCEDIMENTO NUMERICO-DIRETTO: VOLUMI DI TRAFFICO DELLE TRE STAZIONI TIPO  
DIRECT NUMERICAL PROCEDURE TRAFFIC FLOWS OF THREE STATIONS TYPE

Stazione A <i>Station A</i>	Stazione B <i>Station B</i>	Stazione C <i>Station C</i>
N = 159 treni/giorno <i>N = 159 trains/day</i>	N = 122 treni/giorno <i>N = 122 trains/day</i>	N = 67 treni/giorno <i>N = 67 trains/day</i>

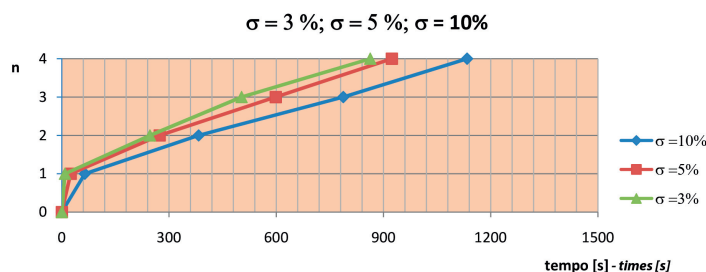


Fig. 8 – Procedimento numerico-diretto: curva di riempimento della Stazione A (modello “con ripetizione”).  
Fig. 8 – Direct numerical procedure: filling curve of Station A (repetition model).

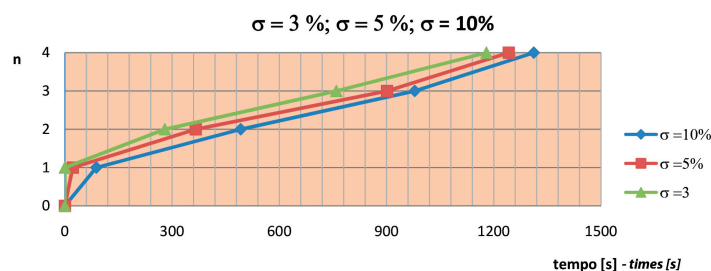


Fig. 9 – Procedimento numerico-diretto: curva di riempimento della Stazione B (modello “con ripetizione”).  
Fig. 9 – Direct numerical procedure: filling curve of Station B (repetition model).

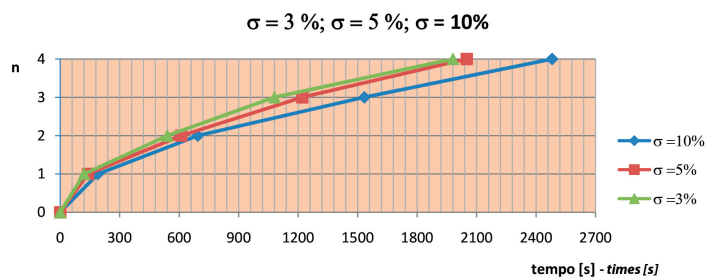


Fig. 10 – Procedimento numerico-diretto: curva di riempimento della Stazione C (modello “con ripetizione”).  
Fig. 10 – Direct numerical procedure: filling curve of Station C (repetition model).

Dalle figure 8, 9 e 10 e dalle corrispondenti tabelle 3, 4 e 5 si possono trarre alcune considerazioni. Innanzitutto si rileva che all'aumentare del valore del parametro  $\sigma$ , a parità del numero di binari, aumenta il tempo  $t_n$  disponibile: infatti accettare un livello inferiore in termini di qualità permette, a parità della dotazione infrastrutturale, di “sfruttare” maggiormente l'impianto. Viceversa, ra-

Some considerations can be drawn from figures 8, 9 and 10 and the corresponding tables 3, 4 and 5. Firstly it should be noted that the increase of  $\sigma$  parameter value, at equal number of tracks, increases the  $t_n$  time available: in fact, accepting a lower level in terms of quality allows to better “exploit” the system, at equal infrastructural level. Conversely, thinking with equal time, accepting a worse service quality allows “serving” more trains.

From tables 6, 7 and 8 the downward trend in the increase of the value of time  $t_n$  as  $n$  increases is also shown; this means that differences between the curves corresponding to different values of  $\sigma$  (however operationally significant) tend to decrease gradually as  $n$  increases (the case of large stations).

A second consideration arising from the analysis is the relationship, with equal  $n$  and  $\sigma$ , between the value of time  $t_n$  and the overall  $N$  number of trains during the reference period; in particular, it can be noted that  $t_n$  decreases as  $N$  increases with an asymptotic trend.

Fig. 11 shows the case for  $n = 4$  and  $\sigma = 10\%$ .

### 3.2. Analysis of split times between train arrivals

From the analysis of timetable data it is possible to find which one better fits the observed data, assuming some probability distributions (normal, exponential, gamma), through the  $\chi^2$  test with a default significance level (table 9).

The test results of  $\chi^2$  for the three stations are shown below and, in the case of Station A, for example, also a graphical representation is proposed.

#### a) Station A

The statistical parameters of the full set of observed frequencies are:

- average = 9.06 minutes;
- variance = 54.04 minutes<sup>2</sup>;

$\chi^2$  test with significance level of 95%, provides the following results for the three functions compared:

TABELLA 3 – TABLE 3

PROCEDIMENTO NUMERICO-DIRETTO:  
VALORI DELLA STAZIONE A  
DIRECT NUMERICAL PROCEDURE:  
STATION A VALUES

Stazione A (tempo $t_n$ espresso in secondi) Station A ( $t_n$ time in seconds)			
	$\sigma = 10\%$	$\sigma = 5\%$	$\sigma = 3\%$
n = 1	65	25	9
n = 2	383	274	247
n = 3	788	599	503
n = 4	1134	923	863

TABELLA 4 – TABLE 4

PROCEDIMENTO NUMERICO-DIRETTO:  
VALORI DELLA STAZIONE B  
DIRECT NUMERICAL PROCEDURE:  
STATION B VALUES

Stazione B (tempo $t_n$ espresso in secondi) Station B ( $t_n$ time in seconds)			
	$\sigma = 10\%$	$\sigma = 5\%$	$\sigma = 3\%$
n = 1	88	22	0
n = 2	492	366	280
n = 3	979	902	760
n = 4	1312	1242	1180

TABELLA 5 – TABLE 5

PROCEDIMENTO NUMERICO-DIRETTO:  
VALORI DELLA STAZIONE C  
DIRECT NUMERICAL PROCEDURE:  
STATION C VALUES

Stazione C (tempo $t_n$ espresso in secondi) Station C ( $t_n$ time in seconds)			
	$\sigma = 10\%$	$\sigma = 5\%$	$\sigma = 3\%$
n = 1	191	140	120
n = 2	694	611	541
n = 3	1534	1221	1081
n = 4	2481	2051	1981

gionando a parità di tempo, accettare una qualità del servizio peggiore consente di “servire” più treni.

Dalle tabelle 6, 7 e 8 emerge inoltre la tendenza alla diminuzione dell'incremento del valore del tempo  $t_n$  all'aumentare di  $n$ ; ciò significa che differenze fra le curve corrispondenti ai diversi valori di  $\sigma$  (comunque operativamente significative) tendono a ridursi man mano che  $n$  aumenta (caso delle grandi stazioni).

TABELLA 6 – TABLE 6

PROCEDIMENTO NUMERICO-DIRETTO:  
CONFRONTO FRA LE CURVE DI RIEMPIMENTO  
DELLA STAZIONE A  
DIRECT NUMERICAL PROCEDURE: COMPARISON  
BETWEEN THE FILLING CURVES OF STATION A

Stazione A - Station A			
	Aumento % di $t_n$ % increase of $t_n$ $\sigma = 10\%$ $\sigma = 3\%$	Aumento % di $t_n$ % increase of $t_n$ $\sigma = 10\%$ $\sigma = 5\%$	Aumento % di $t_n$ % increase of $t_n$ $\sigma = 5\%$ $\sigma = 3\%$
n = 2	56%	40%	11%
n = 3	57%	32%	19%
n = 4	31%	23%	7%

TABELLA 7 – TABLE 7

PROCEDIMENTO NUMERICO-DIRETTO:  
CONFRONTO FRA LE CURVE DI RIEMPIMENTO  
DELLA STAZIONE B  
DIRECT NUMERICAL PROCEDURE: COMPARISON  
BETWEEN THE FILLING CURVES OF STATION B

Stazione B - Station B			
	Aumento % di $t_n$ % increase of $t_n$ $\sigma = 10\%$ $\sigma = 3\%$	Aumento % di $t_n$ % increase of $t_n$ $\sigma = 10\%$ $\sigma = 5\%$	Aumento % di $t_n$ % increase of $t_n$ $\sigma = 5\%$ $\sigma = 3\%$
n = 2	76%	34%	30%
n = 3	29%	9%	19%
n = 4	11%	6%	5%

TABELLA 8 – TABLE 8

PROCEDIMENTO NUMERICO-DIRETTO:  
CONFRONTO FRA LE CURVE DI RIEMPIMENTO  
DELLA STAZIONE C  
DIRECT NUMERICAL PROCEDURE: COMPARISON  
BETWEEN THE FILLING CURVES OF STATION C

Stazione C - Station C			
	Aumento % di $t_n$ % increase of $t_n$ $\sigma = 10\%$ $\sigma = 3\%$	Aumento % di $t_n$ % increase of $t_n$ $\sigma = 10\%$ $\sigma = 5\%$	Aumento % di $t_n$ % increase of $t_n$ $\sigma = 5\%$ $\sigma = 3\%$
n = 2	28%	14%	13%
n = 3	42%	26%	13%
n = 4	25%	21%	3%

Probability function	Degrees of freedom	$\chi^2_{0,95}$
Exponential	1	45,33
Gamma	2	27,90
Normal	2	59,70

Fig. 12 shows trends in observed data (frequency) and theoretical ones, related to the exponential distribution and gamma distribution.

Una seconda considerazione che scaturisce dall'analisi è la relazione, a parità di  $n$  e di  $\sigma$ , fra il valore del tempo  $t_n$  e il numero  $N$  complessivo di treni nel periodo di riferimento; in particolare si rileva che  $t_n$  diminuisce all'aumentare di  $N$  con un andamento asintotico.

In fig. 11 è riportato il caso per  $n = 4$  e  $\sigma = 10\%$ .

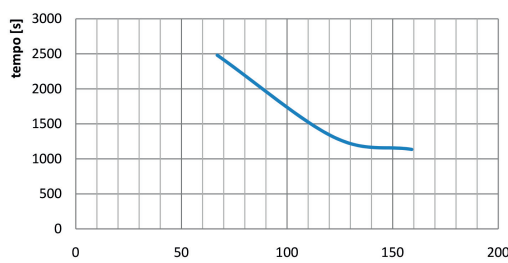


Fig. 11 – Procedimento numerico-diretto: relazione fra tempo  $t_n$  e numero di treno totali  $N$  nel periodo di riferimento.

Fig. 11 – Direct numerical procedure: relationship between time  $t_n$  and total train number  $N$  in the reference period.

TABELLA 9 – TABLE 9

#### ANALISI DEGLI INTERTEMPI: FUNZIONI DI PROBABILITÀ TEORICHE ANALYSIS OF SPLIT TIMES: THEORETICAL PROBABILITY FUNCTIONS

Funzione Function	Densità di probabilità Probability density	Media Average	Varianza Variance
Esponenziale Exponential	$p(x) = \lambda e^{-\lambda x}$	$\lambda^{-1}$	$\lambda^{-2}$
Gamma Gamma	$p(x) = \begin{cases} \frac{\lambda^\alpha}{\Gamma(\alpha)} x^{\alpha-1} e^{-\lambda x} & \text{per } x > 0 \\ 0 & \text{per } x = 0 \end{cases}$ dove $\Gamma(\alpha) = \int_0^\infty t^{\alpha-1} e^{-t} dt$ , con $\alpha > 0$ ;	$\alpha\lambda$	$\alpha\lambda^2$
Normale Normal	$p(x) = \frac{1}{\sqrt{\text{var}(x)}} \frac{1}{\sqrt{2\pi}} e^{-\frac{(x-E(x))^2}{2\text{var}(x)}}$	$E(x)$	$\text{var}(x)$

Dall'analisi dei dati di orario è possibile, ipotizzando alcune distribuzioni di probabilità (normale, esponenziale, gamma), trovare quelle di esse approssima meglio i dati osservati, attraverso il test del  $\chi^2$  condotto con un livello di significatività predefinito (tabella 9).

Di seguito si riportano i risultati del test del  $\chi^2$  per le tre stazioni e, nel caso della stazione A, a titolo di esempio, anche una rappresentazione grafica.

##### a) Stazione A

I parametri statistici dell'insieme completo delle frequenze osservate sono:

- media = 9,06 minuti;
- varianza = 54,04 minuti<sup>2</sup>;

Il test  $\chi^2$  con livello di significatività del 95% fornisce i seguenti risultati per le tre funzioni poste a confronto:

Funzione di probabilità	Gradi di libertà	$\chi^2_{0,95}$
Esponenziale	1	45,33
Gamma	2	27,29
Normale	2	59,70

In fig. 12 è riportato l'andamento dei dati osservati (frequenza) e di quelli teorici, relativi alla distribuzione esponenziale e alla distribuzione gamma.

##### b) Stazione B

I parametri statistici dell'insieme completo delle frequenze osservate sono:

- media = 11,80 minuti;
- varianza = 412,99 minuti<sup>2</sup>;

Il test  $\chi^2$  con livello di significatività del 95%, fornisce i seguenti risultati per le tre funzioni poste a confronto:

##### b) Station B

The statistical parameters of the full set of observed frequencies are:

- average = 11.80 minutes;
- variance = 412.99 minutes<sup>2</sup>;

$\chi^2$  test with significance level of 95%, provides the following results for the three functions compared:

Probability function	Degrees of freedom	$\chi^2_{0,95}$
Exponential	1	53,04
Gamma	2	39,3
Normal	2	50,55

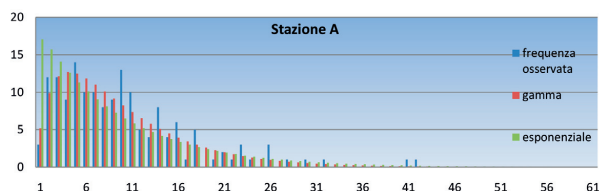


Fig. 12 – Analisi degli intertempi della Stazione A: confronto fra la frequenza osservata e quella teorica derivante da due funzioni di probabilità.  
Fig. 12 – Analysis of split times of Station A: comparison between the observed and theoretical frequency resulting from two probability functions.

Funzione di probabilità	Gradi di libertà	$\chi^2_{0,95}$
Esponenziale	1	53,04
Gamma	2	39,3
Normale	2	50,55

## c) Stazione C

I parametri statistici dell'insieme completo delle frequenze osservate sono:

- media = 21,49 minuti;
- varianza = 442,67 minuti<sup>2</sup>;

Con questa ipotesi il test  $\chi^2$  con livello di significatività del 95%, dà i seguenti risultati per le tre funzioni poste a confronto:

Funzione di probabilità	Gradi di libertà	$\chi^2_{0,95}$
Esponenziale	1	36,19
Gamma	2	25,72
Normale	2	40,45

Si rileva che nei tre casi sia la funzione esponenziale che la funzione gamma garantiscono una buona approssimazione dei dati osservati.

## 3.3. Procedimento teorico statistico

Nelle figg. 13, 14 e 15 e nelle tabelle 10, 11, 12 sono riportate le curve di riempimento calcolate l'approccio numerico iterativo (metodo con ripetizione), rispettivamente per  $\lambda = 2, 4$  e 6 treni/ora, con  $\sigma = 10\%$  e  $d = 8\%$  (errore massimo ammesso).

I valori della variabile  $\lambda$  sono stati scelti in modo da rappresentare situazioni tipiche di traffico delle stazioni ferroviarie.

Nelle tabelle sono indicati anche i valori di  $t_n$  minimi e massimi, associati alle curve di riempimento generate.

TABELLA 10 – TABLE 10

PROCEDIMENTO TEORICO-STATISTICO:  
VALORI DELLA CURVA DI RIEMPIMENTO  
PER  $N=50$  treni/giorno,  $\lambda=2$  treni/ora,  $\sigma=10\%$ ;  
 $d=8\%$

THEORETICAL AND STATISTICAL  
PROCEDURE: FILLING CURVE VALUES FOR  
 $N = 50$  trains/day,  $\lambda = 2$  trains per hour,  
 $\sigma = 10\%$ ;  $d = 8\%$

n	$t_n$ (secondi) $t_n$ (seconds)	Numero iterazioni necessarie Number of iterations needed
0	0	-
1	152	48
2	924	20
3	1937	15
4	3003	11

## c) Station C

The statistical parameters of the full set of observed frequencies are:

- average = 21.49 minutes;
- variance = 442.67 minutes<sup>2</sup>;

With this hypothesis test  $\chi^2$  with significance level of 95%, gives the following results for the three functions compared:

Probability function	Degrees of freedom	$\chi^2_{0,95}$
Exponential	1	36,19
Gamma	2	25,72
Normal	2	40,45

It should be noted that in the three cases both the exponential function and the gamma function guarantee good approximation of the observed data.

## 3.3. Theoretical-statistical procedure

In figures 13, 14 and 15 and tables 10, 11, 12 the calculated filling curves, the numerical iterative approach (with repetition), respectively for  $\lambda = 2, 4$  and 6 trains per hour, with  $\sigma = 10\%$  and  $d = 8\%$  (maximum permissible error) are shown.

The values of variable  $\lambda$  were chosen to represent typical traffic situations in railway stations.

The  $t_n$  minimum and maximum values are also shown in the tables associated with the filling curves generated.

## 3.4. Theoretical-synthetic procedure

Table 13 summarises the results of applying the theoretical and synthetic procedure, assuming  $\sigma = 10\%$  and Poisson arrivals model.

## 3.5. Comparison between results

Table 14 summarises the results of the three procedures.

The results of the numerical-iterative procedure and the synthetic one, for  $n$  significant values ( $n>1$ ), are practically the same; some deviation, however meaningless, can instead be found with the results of numerical-simple procedure.

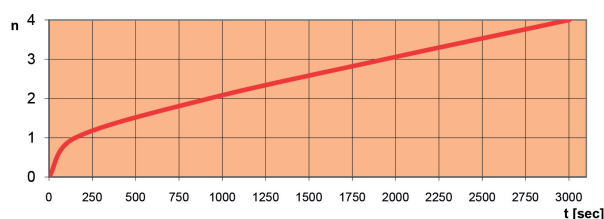


Fig. 13 – Procedimento teorico-statistico: curva di riempimento per  $N = 50$  treni/giorno,  $\lambda = 2$  treni/ora,  $\sigma = 10\%$ ;  $d = 8\%$ .

Fig. 13 – Theoretical-statistical procedure: filling curve for  $N = 50$  trains/day,  $\lambda = 2$  trains per hour,  $\sigma = 10\%$ ;  $d = 8\%$ .

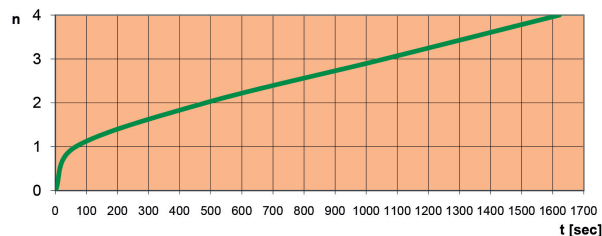


Fig. 14 – Procedimento teorico-statistico: curva di riempimento per  $N = 100$  treni/giorno,  $\lambda = 4$  treni/ora,  $\sigma = 10\%$ ;  $d = 8\%$ .  
Fig. 14 – Theoretical-statistical procedure: filling curve for  $N = 100$  trains/day,  $\lambda = 4$  trains per hour,  $\sigma = 10\%$ ;  $d = 8\%$ .

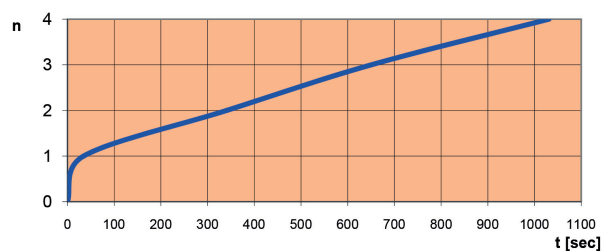


Fig. 15 – Procedimento teorico-statistico: curva di riempimento per  $N = 150$  treni/giorno,  $\lambda = 6$  treni/ora,  $\sigma = 10\%$ ;  $d = 8\%$ .  
Fig. 15 – Theoretical-statistical procedure: filling curve for  $N = 150$  trains/day,  $\lambda = 6$  trains per hour,  $\sigma = 10\%$ ;  $d = 8\%$ .

### 3.4. Procedimento teorico sintetico

I risultati dell'applicazione del procedimento teorico sintetico, assumendo  $\sigma = 10\%$  e modello di arrivi poissoniano, sono riassunti nella tabella 13.

### 3.5. Confronto fra i risultati

Nella tabella 14 sono sintetizzati i risultati dei tre procedimenti.

I risultati del procedimento numerico-iterativo e di quello sintetico sono, per valori di  $n$  significativi ( $n > 1$ ), praticamente uguali; qualche scostamento, tuttavia non significativo, si riscontra invece con i risultati del procedimento numerico-semplificato.

## 4. Conclusioni

Sono stati presentati tre metodi di generazione di famiglie di curve di riempimento delle stazioni ferroviarie.

TABELLA 11 – TABLE 11

PROCEDIMENTO TEORICO-STATISTICO:  
VALORI DELLA CURVA DI RIEMPIMENTO  
PER  $N = 100$  treni/giorno,  $\lambda = 4$  treni/ora,  
 $\sigma = 10\%$ ;  $d = 8\%$   
THEORETICAL AND STATISTICAL  
PROCEDURE: FILLING CURVE VALUES FOR  
 $N = 100$  trains/day,  $\lambda = 4$  trains per hour,  
 $\sigma = 10\%$ ;  $d = 8\%$

n	$t_n$ (secondi) $t_n$ (seconds)	Numero iterazioni necessarie Number of iterations needed
0	0	-
1	65	42
2	483	8
3	1059	9
4	1622	15

TABELLA 12 – TABLE 12

PROCEDIMENTO TEORICO-STATISTICO:  
VALORI DELLA CURVA DI RIEMPIMENTO  
PER  $N = 150$  treni/giorno,  $\lambda = 6$  treni/ora,  
 $\sigma = 10\%$ ;  $d = 8\%$   
THEORETICAL AND STATISTICAL  
PROCEDURE: FILLING CURVE VALUES FOR  
 $N = 150$  trains/day,  $\lambda = 6$  trains per hour,  
 $\sigma = 10\%$ ;  $d = 8\%$

n	$t_n$ (secondi) $t_n$ (seconds)	Numero iterazioni necessarie Number of iterations needed
0	0	-
1	35	26
2	341	3
3	651	4
4	1031	4

TABELLA 13 – TABLE 13

PROCEDIMENTO TEORICO-SINTETICO:  
COORDINATE DELLE CURVE DI RIEMPIMENTO  
THEORETICAL-SYNTHETIC PROCEDURE:  
FILLING CURVE COORDINATES

Volume di traffico Traffic Volume $N$ [treni/giorno]	n	$t_n$ [s]
50	1	182
	2	920
	3	1910
	4	3100
100	1	91
	2	460
	3	960
	4	1600
150	1	61
	2	310
	3	640
	4	1050

TABELLA 14 – TABLE 14

CONFRONTO FRA I RISULTATI  
COMPARISON BETWEEN RESULTS

n	Procedimento teorico sintetico <i>Theoretical-synthetic procedure</i>		Procedimento teorico statistico <i>Theoretical-statistical procedure</i>		Procedimento numerico diretto <i>Direct numerical procedure</i>	
	Volume di traffico (treni/giorno) <i>Volume of traffic (trains/day)</i>	$t_n$ (sec)	Volume di traffico (treni/giorno) <i>Volume of traffic (trains/day)</i>	$t_n$ (sec)	Volume di traffico (treni/giorno) <i>Volume of traffic (trains/day)</i>	$t_n$ (sec)
1	50	182	50	152	67	191
2	50	920	50	924	67	694
3	50	1910	50	1937	67	1534
4	50	3100	50	3003	67	2481
1	100	91	100	65	122	88
2	100	460	100	483	122	492
3	100	960	100	1059	122	979
4	100	1600	100	1622	122	1312
1	150	61	150	35	122	65
2	150	310	150	341	122	383
3	150	640	150	651	122	788
4	150	1050	150	1031	122	1134

In particolare il metodo teorico statistico, generalizzando i risultati del metodo numerico diretto, è uno strumento adattabile a qualunque modello di arrivi e consente di ricavare le curve di riempimento di qualunque stazione ferroviaria, fissata la legge di distribuzione degli arrivi e il volume di traffico giornaliero. Nel caso di studio sono stati compilati tre abacchi riferiti a tre diversi valori del volume di traffico giornaliero (scelti in modo da rappresentare situazioni tipiche di traffico ferroviario) e basati sul modello di arrivi detto di *Poisson*.

Il modello teorico-sintetico, che consente di trovare le soluzioni teoriche dell'equazione integrale (1), è stato sviluppato solo nel caso del modello di arrivi di *Poisson*; la sua generalizzazione richiede ulteriori approfondimenti a causa della complessità dell'espressione matematica che si ottiene adottando leggi di distribuzione degli arrivi più complesse.

L'applicazione dei metodi a casi specifici ha mostrato la sostanziale concordanza dei risultati.

Si ritiene che i risultati ottenuti possano costituire, pur con le necessarie ulteriori conferme sperimentali, validi strumenti nella progettazione e nella gestione dei piani di stazione, al fine di ottenerne una efficiente utilizzazione dal punto di vista tecnico-economico.

#### 4 Conclusions

*Three generation methods of filling curve families of railway stations have been described.*

*In particular the theoretical-statistical method, generalising the results of the direct numerical method, is a tool that is adaptable to any model of arrivals and allows deriving filling curves of any railway station, having set the distribution law of arrivals and the daily traffic volume. In the case study three abacuses referred to three different values of daily traffic volume were made out (chosen to represent typical rail traffic situations) and based on the so-called Poisson arrival model.*

*The theoretical-synthetic model that allows finding the theoretical solutions of integral equation (1), was developed only in the case of Poisson arrivals model; its generalisation requires further study due to the complexity of the mathematical expression which is obtained by adopting more complex arrival distribution laws.*

*The application of the methods to specific cases showed substantial concordance of the results.*

*It is believed that the results obtained may be, albeit with the necessary additional experimental confirmations, valid instruments in designing and managing station plans in order to obtain their efficient use from a technical and economical point of view.*

## BIBLIOGRAFIA – REFERENCES

- [1] CORAZZA – ROTA, "Osservazioni sull'impiego della simulazione per lo studio di nodi ferroviari complessi", Ingegneria Ferroviaria giugno 1992.
- [2] MALAVASI – RICCI, "Affidabilità dei sistemi ferroviari: analisi e valutazioni mediante modelli di simulazione", Ingegneria Ferroviaria gennaio-febbraio 2000.
- [3] CROTTI – DEFLORIO – ELIA – IONCOLI, "Analisi delle potenzialità residua di un impianto di stazione mediante simulazione dell'utilizzo", Ingegneria Ferroviaria luglio-agosto 2005.
- [4] MALAVASI – MARINI – PETRILLI, "Interruzione di esercizio per avaria. Valutazione mediante simulazione", Ingegneria Ferroviaria aprile 1992.
- [5] GIOVINE – LEONARDI, "Un criterio per il dimensionamento di impianti dinamici di manutenzione e pulizia dei rotabili ferroviari", Ingegneria Ferroviaria febbraio 1992.
- [6] DE VITA – GIOVINE – RICCI, "Analisi delle prestazioni di una stazione: Tarvisio Boscoverde", Ingegneria Ferroviaria agosto 2002.
- [7] SPIEGEL, SHILLER, SRINIVASAN, "Probabilità e statistica", McGraw-Hill, 2000.
- [8] CORAZZA – MUSSO, "Impianti di stazione e loro analisi topologica", Ingegneria Ferroviaria novembre 1987.
- [9] CORAZZA – MUSSO, "Circolazione e impianti ferroviari. Verifica a lungo termine", Ingegneria Ferroviaria ottobre 1991.
- [10] CORAZZA – MUSSO, "Circolazione ferroviaria e impianti di trasporto. Stato dell'arte", Ingegneria Ferroviaria luglio-agosto 1991.
- [11] MALAVASI – RICCI, "Analisi di un impianto di stazione sulla base del programma di esercizio", Ingegneria Ferroviaria ottobre 1989.
- [12] CORAZZA – FLORIO, "Il problema del nodo e la verifica degli impianti di stazione", Ingegneria Ferroviaria aprile 1979.
- [13] FLORIO – MALAVASI – SALVINI, "Analisi dei ritardi e verifica di un nodo ferroviario complesso", Ingegneria Ferroviaria luglio 1985.

## Sommaire

GÉNÉRATION DES COURBES DE REMPLISSAGE  
DES GARES DE TRAIN.

La méthode nommée "des courbes de remplissage" permet, avec une approche synthétique, ça veut dire ne pas lié à la topologie de la gare mais seulement à certains paramètres "globaux" comme le nombre de rails, le projet ou la vérification de la capacité de stationnement des gares de train, associée à un niveau de qualité de service.

Le but de ce mémoire est de décrire certains méthodes pour la construction de courbes de remplissage des gares de train, en fonction pas seulement du niveau attendu de qualité, mais aussi du volume de trafic et de la loi de distribution des arrivées.

## Zusammenfassung

BERECHNUNG DER FÜLLUNGSKURVEN VON  
EISENBAHNBÄHNHÖFEN

Die Verfahren, die auf die Füllungskurven gestützt sind, erlauben die Ermittlung der Gleiszahl unabhängig von der Topologie der Gleisverknüpfung. Dazu aber können andere Faktoren in Kauf genommen werden, wie z.B. das Qualitätsniveau des Betriebs. Es werden verschiedene Methoden vorgestellt, die die rechnerische Zeichnung der Füllungskurven für Eisenbahn Bauhöfen gestatten, in denen auch die Zugezahl in einem bestimmten Zeitintervall und die statistische Ankunftsverteilungen als Parametern gelten.