# How regular is a regular-interval timetable? An operational tool to assess regularity 

David Tron, EPFL-LITEP<br>Panos Tzieropoulos, EPFL-LITEP

Conference paper STRC 2009

9 $^{\text {th }}$ Swiss Transport Research Conference
Monte Verità / Ascona, September 9. - 11. 2009

# How regular is a regular-interval timetable? An operational tool to assess regularity 

David Tron, EPFL-LITEP<br>EPFL-LITEP<br>Lausanne

Phone: +41 216932479
Fax:
email:
secretariat.litep@epfl.ch

Dr. Panos Tzieropoulos<br>EPFL-LITEP<br>Lausanne

| Phone: +41216932479 | Phone: |
| :--- | :--- |
| Fax: | Fax: |
| email: | email: |

August 2009


#### Abstract

Assessing the regularity of a timetable is a step-by-step procedure. At the initial stage, there is a need to set several options that are specific for every single case, and to state the problem in a systematic, structured way. Transport supply, for instance, should be defined as a set of specific train services ("missions") that not only differ in travel time or dwell policy, but also reflect the operator's commercial strategy. Some of the settings can be defined without ambiguity; other ones are not suitable for a broad and exclusive definition and must therefore be regarded in a case-by-case basis (e.g. the operational range, i.e. the time span of train operations that has to be considered for the evaluation). This necessity of a case-by-case handling has been a major source of complexity in developing software for assessing the regularity of a timetable.

On the other hand, automation of the assessment process makes it possible to treat in detail significant amounts of data that would be impossible to handle in a manual assessment. By virtue of automation, it is possible to set the geographical focus and to zoom-in in order to compare and analyse not only the main O/D links, but parts of them as well. It is also possible to use multiple sets for the initial parameters, and to vary the tolerances used in measuring regularity. This is particularly helpful if one needs to proceed with sensitivity analyses. Finally, automating the assessment process offers great versatility in displaying the results according to the user's wishes.

The automation requires three inputs from the user. The first input is the service backbone, defined as a set of standard train paths. This is the reference to which the actual timetable will be compared to. The second input is the actual timetable. The third input are the tolerances upon which computed differences may qualify as significant.

The paper describes the software developed to automate the assessment process based on a methodology, which is presented in another paper (by Tzieropoulos and Emery). Then, it comments the main difficulties and achievements in developing the software. Finally, it shows the utility of the developed tool, its possible applications, and the advantages resulting from the automation of the assessment process.


## Keywords

Regular-interval timetable - coordinated regular interval timetable - timetable regularity

## 1. Introduction

This paper relies on the regular timetable measure method as described in the paper by Tzieropoulos \& Emery "How regular is a regular-interval timetable? Theoretical foundations and assessments methodology" $[7]^{1}$. It takes over the same definitions and focuses on how the developed methodology has been implemented in an operational tool to help policy makers assessing the regularity of actual timetables. The main elements of the methodology are:

- The assessment is made by comparing a pre-defined reticular diagram, which is the graphical representation of the basic timetable framework. The reticular diagram displays the services, their periodicity (the time interval at which the train paths are provided) and the departure and arrival minutes in major stations.
- The regularity of the timetable is computed for a given operational range (the time span of train operations that has to be considered for the assessment).
- 3 indexes are computed:
o A structure index, reflecting how well the different services comply with the service backbone
o A regularity index, reflecting how well the timetable complies with the periodicity defined in the basic framework
o A reinforcement rate, reflecting the volume of extra train paths aimed to enhance peak period services
- Tolerances are included while assessing compliance of train paths with both the service backbone and the regularity. Strict compliance means that there is no difference with the basic framework. In case of small differences that lay within the tolerance thresholds, compliance is qualified as loose. If a given train path transgresses any tolerance rule, it is qualified as outlier.

The developed methodology has been firstly applied manually to assess structure and regularity of the 2008 timetable for the French Rhône-Alpes Region. Eventually, the process has been automated to be able to deal with larger input volumes to provide better flexibility in terms of possible results. The scope of the paper is to present the automation of the process.

[^0]
## 2. Limits of the manual approach

Applying the methodology in a real-life case quickly pointed out some limits of the manual approach.

### 2.1 Assessing train paths to services

Every single train path of the timetable must be compared in terms of departure/arrival minutes, travel time, and intermediate stops with the defined "service backbone" in order to associate it to the right service. In complex cases where the evaluation of a whole network is needed, this "train sorting out" must be done for up to several thousand trains. Needless to say that that a substantial amount of time is needed for this operation.

### 2.2 Changing the parameters and flexibility

Tolerance values used to qualify compliance as strict or loose are neither unique nor fixed. These values can and must be discussed depending on the particular network and location as well as the objectives of the assessing body. The same can be said regarding the operational range.

If after a first evaluation of the regularity any of these parameters needed to be changed, the only option left with a manual approach would be to restart the whole procedure from the beginning. This is extremely time consuming.

### 2.3 Accuracy of the results

To be practical, the manual approach takes only into account time differences at the origin and destination of the train paths. For some diametric connections (that run across major stations) or for services having, e.g., a 60 minutes time interval on their main section and 120 minutes on end-of-line extensions, service variations in intermediate stations should also be taken into account. As soon as the number of intermediate stops that need to be counted rises, manual assessment becomes extremely cumbersome. For instance, for a train going from A to D through B and C , taking into the account the intermediate stops means that one has to evaluate 6 O/D links: A-B, A-C, A-D, B-C, B-D and C-D.

The more the manual process becomes complex, the more the probability of human errors increases. Those reasons point out the necessity to develop a tool that fastens and widens the possibilities of all or part of the assessment process.

## 3. Concept of the algorithm

Basically, the algorithm does nothing else than doing automatically what was previously done manually: compare a train path to a "standard service" and then determine whether it is fully compliant, mostly compliant or essentially different.

However, due to the complexity of a timetable and the great amount of information that the comparison produces, what sounds like a simple comparison algorithm proved to be more complicated when put into practice.

### 3.1 Inputs

Three different input categories can be distinguished:

- The timetable
- The basic timetable framework
- The evaluation parameters


### 3.1.1 The timetable

This input represents the information that has to be evaluated. For every train path in the timetable, the following information should be provided:

- A train number that defines the train path in a unique way
- The station list that includes the origin, the destination, and every station the train stops in
- For each station in this list, an arrival and/or a departure time ${ }^{2}$
- The operating days or period, to be able to exclude from the assessment trains that run only once or under particular conditions (This input might be skipped if the timetable is already containing only the trains running on standard weekdays)

Input of timetable information should be fully automatic, assuming a readable dataset is available. Raw data may need some reformatting but this can be easily automated.

[^1]Figure 1: Timetable input sample


### 3.1.2 The basic timetable framework

The basic timetable framework covers a 60 - or 120 -minutes of operation. That pattern is meant to be repeated identically during the full operational range and that leads to a $100 \%$ regular timetable.

For each service included in the basic framework, the following information must be provided:

- Type of service (e.g.: IC, IR, R, etc.)
- The list of serviced stations
- Arrival and/or departure minute for every station of the list
- The time interval between successive train paths belonging to the same service (e.g. 30 minutes, 60 minutes, etc.)

The user must be in charge of defining the basic timetable. In most case of actual timetables, the information on the basic framework is already available and needs only to be input to the tool. In some cases, the basic framework must be manually revere-engineered based on the actual timetable. The developing team made the deliberated decision to not automate this step of reverse engineering. Very often, second guessing an underlying basic framework from an actual timetable leads to many possible alternatives. However, the basic framework itself is
the technical translation of a transport policy set according to service and commercial criteria. Moreover, the same timetable assessed against different basic framework produces significantly different outcomes. It is therefore important that, through explicit definition of the basic framework, the user sets the policy vision against which the time table will be assessed.

Figure 2: Format of a basic framework input sample

|  | Départure/arrival | Mission Type | Time interval |  | Départure/arrival | Mission Type | Time interval |
| :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: |
| Genève | 3 | R | 60 | Allaman | 7 | R | 60 |
| Genève-Sécheron | 5 |  |  | Etoy | 9 |  |  |
| Chambésy | 8 |  |  | St-Prex | 11 |  |  |
| Les Tuileries | 9 |  |  | Tolochenaz | 14 |  |  |
| Genthod-Bellevue | 11 |  |  | Morges | 18 |  |  |
| Creux-de-Genthod | 14 |  |  | Renens VD | 26 |  |  |
| Versoix | 16 |  |  | Lausanne | 31 |  |  |
| Pont-Céard | 17 |  |  |  |  |  |  |
| Mies | 20 |  |  | Genève | 21 | RE | 60 |
| Tannay | 21 |  |  | Coppet | 31 |  |  |
| Coppet | 26 |  |  | Nyon | 36 |  |  |
|  |  |  |  | Nyon | 37 |  |  |
|  |  |  |  | Gland | 42 |  |  |
| Genève | 45 | IC | 60 | Rolle | 48 |  |  |
| Lausanne | 17 |  |  | Allaman | 53 |  |  |
|  |  |  |  | Renens VD | 7 |  |  |
|  |  |  |  | Lausanne | 12 |  |  |

### 3.1.3 Evaluation parameters

The following parameters should be set by the user (Figure 3):

- Minimum operational range that may also be expressed as the minimum number of train paths expected in order to achieve $100 \%$ regularity.
- Tolerance range for the departure minute (low boundary for early departure, and high boundary for late departure)
- Tolerance range for the arrival minute (low boundary for early arrival, and high boundary for late departure)
- Maximum difference allowed on the travel time (a single value, meaning + or - "x minutes" allowed compared to the standard time)
- Maximum difference in stopping policy (a single value, as the number of less or more stops allowed compared to the standard service)
- Tolerance threshold for determining reinforcement trains planned at half the regular interval (a single value, e.g.: giving a tolerance of 5 minutes means than peak-period reinforcement trains can be located, assuming a regular interval of 60 minutes, 25 to 35 minutes after the regular train)

This last parameter is an add-on to the manual approach. It aims to more accurately determine the reinforcement rate. The only train paths qualifying as reinforcement trains are those running at half the regular time interval.

There is still an open debate of applying the algorithm twice: once for peak periods and once for off-peak. By setting the last threshold, we can set aside this debate. A large percentage value of the reinforcement rate may suggest revisiting the definition of the service by setting the regular interval at half the initial value.

Figure 3: Inputting tolerance thresholds


### 3.2 The evaluation process

### 3.2.1 Data

The service backbone is displayed as a set of services sorted by type and O/D. Travel times and the regular interval associated to each service are displayed as well (Figure 4).There is also an option to display the full details of each service (Figure 5).

Assignment of train paths to services is done during the introduction of timetable data ${ }^{3}$. Trains paths are immediately associated to a standard service if possible. They appear sorted by service, train number, O/D and travel time (Figure 6).

Detailed results for timetable's train paths are also available (Figure 7); they will be presented in a later section of the paper. This figure also shows that arrival and departure minutes can but must not be provided. For example, in the shown case, only the departure minutes are provided for the intermediate stations, resulting in a stopping time equal to 0 .

[^2]Figure 4: Displaying the service backbone


Figure 5: Details of a standard service


Figure 6 Displaying timetable's train paths

| 56 MesCad |  |  |  |  |  |  |  |  |  |  | - $\square$ |
| :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: |
| Elie Help |  |  |  |  |  |  |  |  |  |  |  |
| Nouveau | Ourrir |  | $\stackrel{\text { mport sillons Types }}{\text { s. }}$ | $\begin{gathered} \text { ©t } \\ \text { Import Hora } \end{gathered}$ |  | $\underset{\text { Tolérances }}{\substack{\text { She }}}$ |  |  |  |  |  |
| Sillons types | Horaire | Gares | Resultats |  |  |  |  |  |  |  |  |
| Catégorie |  |  | Numéro | - | Dépar |  | , | Arrivée | 4 | Temps de parcours |  |
| $\begin{aligned} & \pm \mathrm{IC} \\ & \pm \mathrm{R} \\ & \boxminus \mathrm{RE} \end{aligned}$ |  |  |  |  |  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |  |  |  |  |  |
|  |  |  | 886804 |  | Genè |  |  | Lausanne |  |  | 50 minutes |
|  |  |  | 886807/6 |  | Genè |  |  | Lausanne |  |  | 52 minutes |
|  |  |  | 889957 |  | Genè |  |  | Lausanne |  |  | 51 minutes |
|  |  |  | 887708 |  | Genè |  |  | Lausanne |  |  | 51 minutes |
|  |  |  | 886265/4 |  | Genè |  |  | Lausanne |  |  | 51 minutes |
|  |  |  | 888913 |  | Genè |  |  | Lausanne |  |  | 51 minutes |
|  |  |  | 888917 |  | Genè |  |  | Lausanne |  |  | 51 minutes |
|  |  |  | 889967 |  | Genè |  |  | Lausanne |  |  | 51 minutes |
|  |  |  | 886834/5 |  | Genè |  |  | Lausanne |  |  | 51 minutes |
|  |  |  | 88773 |  | Genè |  |  | Lausanne |  |  | 51 minutes |
|  |  |  | 888929 |  | Genè |  |  | Lausanne |  |  | 51 minutes |
|  |  |  | 88774 |  | Genè |  |  | Lausanne |  |  | 51 minutes |
|  |  |  | 886287/6 |  | Genè |  |  | Lausanne |  |  | 51 minutes |
|  |  |  | 886856/7 |  | Genè |  |  | Lausanne |  |  | 51 minutes |
|  |  |  | 888939 |  | Genè |  |  | Lausanne |  |  | 51 minutes |
|  |  |  | 2349 |  | Genè |  |  | Lausanne |  |  | 51 minutes |

Figure 7: Details of the assessment results for timetable's train paths


### 3.2.2 The stations

During data input, the programme creates automatically a list of all the stations included in the timetable or in the basic timetable framework. Stations are divided in two categories:

- G1 being the stations that are used to assess train path compliance in terms of time scheduling
- G2 being the stations that are only taken into account to verify that the train's stopping pattern matches the standard service's.

By default, only origin and destination of train paths are assigned to type G1; all other intermediate stations are assigned to type G2. However, users may eventually declare of type G1 chosen intermediate stations, to make them count in the assessment. Therefore, on a given O/D there might be several G1 stations and therefore several sections for which results will be available. E.g. when considering the section between Geneva and Lausanne, the following O/D may appear in standard services:

- Geneva - Lausanne
- Morges - Lausanne
- Allaman - Lausanne
- Coppet - Genève

Now, for every service between Geneva and Lausanne, partial evaluation is possible with the stations Morges, Allaman and Coppet, provided that the service has a stop in these stations.

Figure 8: Example of stations listing

| Sillons types | Horaire | Gares | Resultats |
| :--- | :--- | :--- | :--- |
| Nom | Catégorie |  |  |
| Allaman | G1 |  |  |
| Chambésy | G2 |  |  |
| Coppet | G1 |  |  |
| Creux-de-Genthod | G2 |  |  |
| Denges-Echandens | G2 |  |  |
| Etoy | G2 |  |  |
| Genève | G1 |  |  |
| Genève-Sécheron | G2 |  |  |
| Genthod-Bellevue | G2 |  |  |
| Gland | G2 |  |  |
| Lausanne | G1 |  |  |

This example of Figure 8 points some limits of the default station type assignment, attributing type G1 only to the stations that are origin or destination of a service: A station like Nyon, which is more important than stations like Coppet or Allaman but - contrary to them - is neither an origin nor a destination - will not be taken into account automatically while assessing regularity. It is thus, up to the user, to declare of type G1 important intermediate stations. Notice that any difference in a G1 station counts in assessing structure and regularity. Thus, even if the departure and arrival minutes at the origin and destination are identical to the basic framework but there is a small difference at any of the G1 intermediate stations, the train path compliance will qualify as "loose".

This means that adding G1 stations not only increases the number of relations for which results will be available, but makes also the evaluation more sensitive to small differences: train paths that whose compliance would qualify strict otherwise, may be classed as "loosely" respecting the regularity or the structure.

### 3.2.3 Results

Basically, the programme computes the structure index, the regularity index, and shows for any given section the split of train paths among their different types (Figure 9). The latter are:

- Train paths complying strictly with the regular-interval principle
- Train paths complying loosely with the regular-interval principle
- Train paths reinforcing the peak-period service and perfectly complying with the service characteristics
- Train paths reinforcing the peak-period service and complying loosely with the service characteristics
- Train paths perfectly complying with the service characteristics but not scheduled at regular time intervals
- Train paths complying loosely with the service characteristics and not scheduled at regular time intervals
- Outliers

Train paths are labelled following the above shown order. As soon as a train path matches a type, it is labelled. We thus avoid double labelling of train paths.

Figure 9: Sample of results for a given service on a section

## Mission R Genève- Coppet

entre Genève et Coppet

## Taux de cadencement



## Taux de structuration



Total: 18
Structurés stricts: 13 (72.2\%)
Structurés souples: 5 (27.8\%)
Hors système: 0 (0.0\%)

Répartition détaillée


## Fictive example that does not reflect the score of the real SBB timetable

## 4. Advantages of the implemented approach

As already mentioned there are several advantages in automating the assessment process:

- Fast evaluation: Despite the initial effort needed for input data preparation, the evaluation of a whole network is extremely fast, almost instantaneous.
- Parameter changes and sensitivity analyses become possible: Once the first evaluation is done, re-evaluating the timetable with alternative tolerance thresholds is a matter of changing a few values and exporting the new results.
- More accurate and representative results by taking into account intermediate stops: Thanks to the possibility to add any station in the assessment process by declaring it as of type G1, the results are much more accurate than sticking with the regularity in only the origin or the destination of a given service. This also eliminates the problem of the diametric connections. E.g., a Zurich - Geneva train is diametric for several important stations. With intermediate stations set to G1 it is possible to obtain results for every desired link. However, as already mentioned, by so multiplying the volume of G1 stations, there is a probability of increasing the score in loose compliance.
- Easily possible to refine the network modelling by adding intermediate stops later into the process: As for changing of parameters, adding intermediate stations after a first evaluation is easily possible.
- The user still has control on the structural inputs (basic timetable framework): The user still has a great influence on the results as the basic framework is not automatically inferred from the timetable but should be manually defined. This ensures that the assessment process is done using a clearly and explicitly defined target that actually reflects the transport policy against which a given timetable is compared.


## 5. Limits of the approach and the implementation

Although the semi-automatic method has many advantages in comparison to a manual assessment, some not standard situations need an arbitrary decision and are difficult to evaluate automatically and their taking into account still needs to be solved

- Sensitivity on the choices related to the operational range, and namely its duration and positioning. This limit triggers, for instance, when dealing with early morning and late evening outliers: First and last trains, just out of the boundaries of the operational range, are often scheduled with more or less slight deviations compared to the service they belong to. Determine whether a train path has to be included within the operational range is intuitively easy. Doing it automatically proves to be more difficult. One solution is to admit that the lower boundary of the operational range is located at the first train found that matches the regular-interval operation. This solution works well for dense and highly regular services. Applying such a rule in cases of timetables with very few regular trains becomes tricky. In those cases a fixed start and end time for the operational range must be set. It is however difficult to determine unique values valid for all services and O/Ds. As for now, the solution adopted is to start the operational range with the first complying train and to end it with the last one, even though this option may lead to low index values.
- Reinforced peak-hour services are detected only if scheduled at half the regular interval. Theoretically, for a service planned with a 60 -minutes interval, reinforcements during peak periods can be scheduled not only with a 30-minutes interval, but possibly also every 15 minutes. As by now, the programme detects only those planned at half the regular interval, and misses any ones that would have been planned at a quarter (and by that matter at any fraction ${ }^{4}$ ) of the regular interval.
- Two-hour-interval services: Nothing is provided in order to determine if a two-hourinterval service is supposed to be operating on the slice of even or odd hours. What is more, there are very often one-hour-interval reinforcements at peak hours that blur the determination of the hour slice to take into account. This is solved for now by a "majority rule", choosing the hour slice that gives the best score for the regularity index. This choice is not transparent for the user.

[^3]
## 6. Conclusions and work ahead

Applying the assessment methodology in real life cases - first manually, and then thanks to its automated implementation - showed that it produces sound, meaningful, and useful results that actually helped the policy makers ([1], [2], [5]).

Meanwhile, development of the software is not over yet. Many improvements are still on the drawing board and a lot of testing is still necessary to make sure the package is able to deal with any possible configuration of the service backbone (including some weird ones).

Stronger aggregation of the results and their representation is main ongoing improvement. At the moment for instance, it is possible to get results for every section of every service. It is also possible to provide aggregated results of several services on any given O/D, but not yet for any given intermediate track section.

A pending development is to compute regularity and structure indexes for stations, showing for instance the percentage of regular or structured trains that cross a given station. This information should be provided for G1 as well as for G2 stations.

All in all, for high level policy makers, the tool should be able to provide some comprehensive indexes covering the whole network and the full set of services. Doing so, however, raises some tricky aggregations and weighing issues for which it is difficult to provide consistent answers.

## 7. Credits

The authors want to express their gratitude to Joël Schintgen, ing. EPFL, who has been in charge of refining the algorithms and implementing the automated procedure. The current tool has been developed under Python and runs on Windows and Linux platforms. Joël's involvement in the development has been decisive in translating a general methodology to an operational tool. Credits are also due to Daniel Emery (EPFL-LITEP) who proceeded with the first manual application of the methodology.

## 8. References

[1] Émery, Daniel (2009), Mesure du cadencement, Note technique $\mathrm{N}^{\circ}$ 2, Retour d'expérience sur la mise en service du cadencement 2008 en Rhône-Alpes, EPFLLITEP, Lausanne
[2] Perrin, Jean-François (2009), Retour d'expérience sur les cadencements régionaux lancés au service 2009, Thèse professionnelle, Mastère d'action publique, Mission réalisée au sein de Réseau Ferré de France, École Nationale des Ponts et Chaussées, Paris
[3] Rey, George (2007), Entwicklung des ITF von den Anfängen bis zur Gegenwart, Eisenbahntechnisches Kolloquium, Darmstadt, June 5
[4] Stohler, Werner (2003), Why is an integrated clockface-driven railway system more efficient than a divided competition-oriented railway system? SMA und Partner AG, Zürich
[5] Tzieropoulos, Panos et al (2008), Diagnostic, Retour d'expérience sur la mise en service du cadencement 2008 en Rhône-Alpes, EPFL-LITEP, Lausanne
[6] Tzieropoulos, Panos et Émery, Daniel (2009), De la théorie à la pratique, in Préconisations, Retour d'expérience sur la mise en service du cadencement 2008 en Rhône-Alpes, EPFL-LITEP, Lausanne
[7] Tzieropoulos, Panos et Émery, Daniel (2009), How regular is a regular-interval timetable? Theoretical foundations and assessment methodology, also to be presented in STRC 09


[^0]:    ${ }^{1}$ Numbers between brackets point to the reference list at the end of the paper.

[^1]:    ${ }^{2}$ It is not compulsory to input the arrival and departure time for every station. Providing only the departure time is sufficient for the tool to run, with some obvious loss of accuracy.

[^2]:    ${ }^{3}$ The basic framework can also be input after the timetable. Without data on the basic framework, train paths are displayed unsorted.

[^3]:    ${ }^{4}$ The authors are aware of a real life case, where reinforcements are planned at 20 -minutes interval for an hourly service.

