Paper accepted for presentation at
First Interdisciplinary Workshop on ML, Law and Society, track Knowledge Discovery and process mining for Law (KDPM4LAW)
at the ECML-PKDD Conference, Turin, Sept. 22, 2023

Variants analysis in judicial trials: Challenges and initial results

Alessandro Campi[0000−0001−8138−4146], Stefano Ceri[0000−0003−0671−2415], Marco Dilettis, and Barbara Pernici[0000−−0002−−2034−−9774]

Politecnico di Milano - DEIB {nome.cognome}@polimi.it

Abstract. Predictive process mining aims to provide information about the future evolution of processes, given the information available from past process executions and about the current state. In this paper, we focus on judicial trials as a specific type of process and we analyse the aspects that have an impact on process duration and the challenges posed by this type of analysis.
In particular, we focus on analyzing and discussing the factors that have most impact on the duration of trials. An analysis framework is proposed, which takes advantage of a large dataset describing five years of trials in the Court of Appeal of Milan. We examine both the phases and total length of the trials and we propose techniques to identify events that are potentially critical, as they have a major impact on their duration.

Keywords: Process mining · Data mining · judiciary.

1 Introduction

Europe is concerned with the length of trials in Italy, hence several EU-funded projects are focused on the direction of monitoring the performance not only of terminated cases, but also of ongoing trial. The increased use of information systems allowed analysing more and more in depth some Key Performance indicators, such as the Disposition Time (DT) and the Clearance Rate (CR) of judicial trials. While recent data reports show that the backlog is reducing, they still show that Italy has much longer trial duration than other European countries[1]. The duration of trials has a broad impact on the society, including also economic implication [3]. As a result the attention on this theme is high both at the National and European level.

The goal of this paper is to propose a process-driven approach, based both recognizing process states and events that cause state changes at the macro level, taking advantage of new process mining techniques. We present preliminary results based on a large case study.

The paper is structured as follows. In Section 2, we discuss related work. In Section 3, we present the scenario motivating our study, based on data of the

[1] https://www.coe.int/en/web/cepej/special-file-report-european-judicial-systems-cepej-evaluation-report-2022-evaluation-cycle-2020-data-

Court of Appeal of Milan. In Section 4, we discuss different levels of representation of judiciary workflows. In Section 5, we introduce the approach proposed in this paper to analyse existing events and states traces and we discuss the initial results. Finally, in Section 6 we present our concluding remarks.

2 Related work

Process mining [2] allows analysing a process through the steps of data extraction, data preparation, process discovery, conformance and compliance checking, and performance analysis. Process mining techniques and data mining have been studied in recent works [9, 6], not only to analyse the global behaviour of processes, but also to identify outliers and critical situations. Recently, process mining has been augmented by AI-inspired techniques [4]. In [5] waiting times are investigated, classifying different types of causes of delays such as batching or resource contention, and some initial methods to identify them on the basis of mining activity transitions are proposed.

Process mining has been applied to judicial systems in Brazilian courts [9] to derive process maps which are used to identify slow transitions and activity bottlenecks and to analyse the times of processes on the basis of different analysis dimensions, e.g., comparing paper-based and digital processes. In [6], process mining based on causality graphs is performed considering outlier cases, which allows identifying the main events that may delay the processes. However, this type of analysis does not allow to identify causes of delays that are not linked to pairs of events and further research is needed to analyse the impact of events in general.

Recent work is focusing on predicting process execution times [1, 5]. Recent directions in applying AI techniques to trials are presented in [7], where different deep learning techniques are applied to predict the duration of a phase. However, in these analyses, the sequences of events are not being considered and the focus is on a single phase. More general methods are needed, taking into consideration the different possible variants of processes and sequences of events.

In our previous work [8] we started discussing directions to analyse not only sequences of events, but also states and event impacting trial duration. In this paper, we develop our initial ideas further, and propose a new approach, based o a data-driven analysis of states and events, for monitoring critical cases and for predicting the duration of trials.

3 Scenario

Our work is based on a large case study, including all trials started after January 2017 and until March 2023 in the Court of Appeal of Milan, along with the "Ordinary second degree process procedures" . The scenario considered in the paper is supported by a process management system (PCT - Processo Civile Telematico), governed by rules defined for all Courts at the National level. The system defines all possible states and transition events between states; it is used by presidents, judges, and chancellors, assisted by specific interfaces.

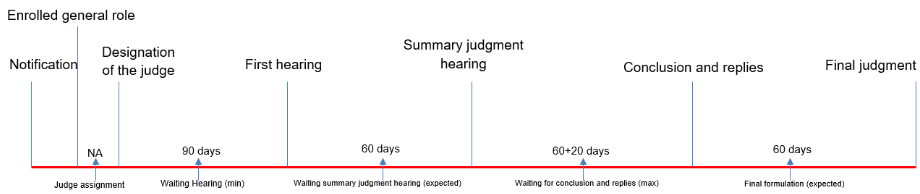

Fig. 1. Main phases in a trial

The main trial phases are represented in Fig. 1. While this diagram represents a standard process flow, a very high number of other possible states and events are defined to represent all possible cases. In fact, even if looking only at second degree - or appeal - trial procedure, 48 possible states for a trial are defined, from the initial states related to Judge Assignment until the state of Final Judgment, including both regular flows of trials as represented in the figure and other states that can arise in a trial according to the laws. Within these states, over 3,600 different events are encoded in the system to represent the complete workflow.

The flow of events is quite regular, following periodicity due to weekends, summer and winter breaks. In the considered period, the COVID pandemic caused a strong reduction of cases arrival and management, in particular in Spring 2020. The number of events recorded in the system is shown in Fig. 2, restricted to completed cases in the period.

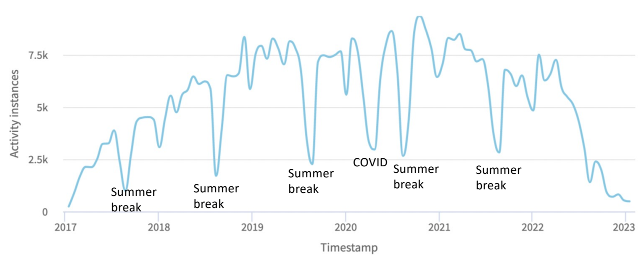

Fig. 2. Number of events for completed trials started between Jan. 2017 and Mar. 2023

This amount of data enable a data-driven monitoring and management of pending cases, identifying and predicting trial duration, delays, and critical events causing delays.

4 Representing judiciary workflows

The workflow of judicial cases can be represented at three different levels, later explained.

4.1 Real world level: document flow, steps, decisions

The data at our disposal represent the states and events that describe what happens to a process during its course. The recording of events (performed by the judge or chancellors, usually corresponding to formal acts and documents) leads the process to change its state.

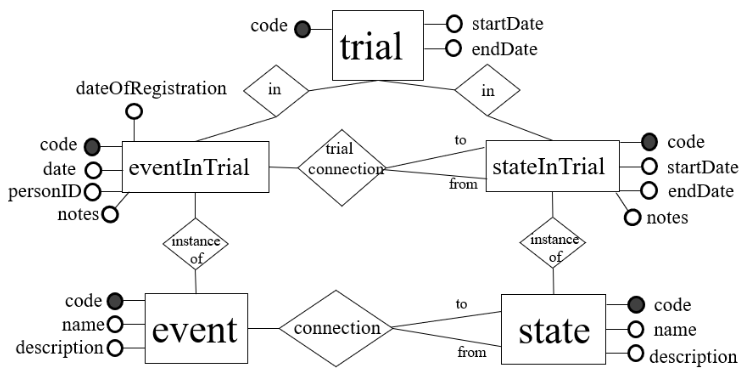

Fig. 3. Schema of the data

4.2 Conceptual level: logical states and events

The Entity-Relationship (ER) diagram in Fig. 3 represents trials. The trial entity is connected to its events and states; these represent the data as they are entered in the process control system. A ternary relationship connects each event, which causes a state transition, with the states respectively before and after the event (a state transition is uniquely identified only by the event id and the ids of the initial state of the transition and of the reached state after the transition). While at this level we represent what indeed happens in the trial, events and states are described at an abstract level, describing what is legally allowed; also in this case, a ternary relationship connects the event to two states, describing a legal state transition. Entities are equipped with descriptive attributes, and each entity and ternary relationship is mapped to a relational table by using standard transformations.

4.3 System level: events registration in the workflow system

The events that have occurred are recorded in the system by judges or chancellors. Some events may be entered immediately (e.g., a judge inserting the minute of a verdict), but in many cases events are recorded later by chancellors, and some delay may occur in the registration, in particular when a paper flow is associated with the event (e.g., an act signed manually or a paper folder for the trial). This inevitably introduces latency between what actually happens and what was stored in the database, leading to inconsistencies. For instance, it often happens that events occurred over a fairly long period of time are recorded all together, perhaps thanks to cascading effects.

4.4 Issues

These three levels correspond to different representations of the same reality, however mismatches are possible for different reasons:

- Transition rules are governed by strict legal procedures and are encoded in the data as described by a finite state automaton that allows all and only the steps provided by the law. Unfortunately, due to legislative changes, the laws and the automaton describing the lawful steps are not always synchronized because there is a time gap between the approval of a law and the subsequent adjustment of the automaton. In this transitional period, transitions or states stated by law are not encoded by the system, and therefore some manual interventions (workarounds) are made, leading to inconsistency in the data.
- Some events are batched, so their occurrence depends on the time window in which they occur (assignment of cases to judges, hearings).
- Event registration can also be batched, resulting in delays of event registration w.r.t. its occurrence.
- Not-recorded events: some real world events might not be recorded in the system to enable the trial continuation, that has indeed occurred (e.g., minutes revisions, paper flows which are constraining the trial evolution).

5 Approach

The ingredients behind our approach are illustrated in Fig. 4. Constraints on possible sequences of states and events and temporal constraints, such as legal liabilities for overtime trials or specific phases, are defined in the rules. A process management system, with subsystems for process enactment and tracing, produces process data and meta data; it is used by judges and chancellors. The focus of our work is to develop reasoning tools to assess ongoing processes in general, and to identify critical events and states that might impact duration of ongoing trials, and that can potentially both rise alerts for critical situations, and also provide a general overview to high-level monitoring subjects, such as presidents and inspectors.

For reasoning about the state of processes, process mining and data analysis techniques are developed for the following goals:

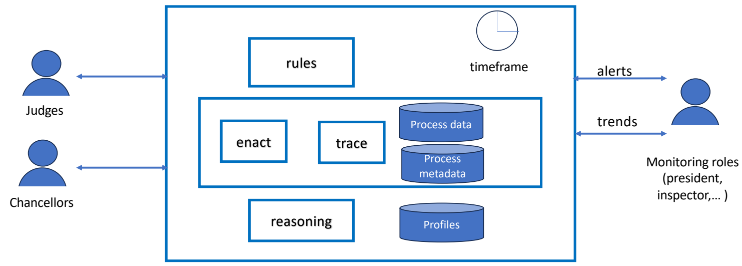

Fig. 4. Process mining for the judiciary

- Variant analysis and their assessment: the variants in process execution are analyzed for monitoring purposes.
- Identification of critical events: the impact of specific events on the length of execution of a process is assessed, to single out events that are systematically associated to anomalous situations.
- Predictive approaches for alerts: the duration of processes and states are analyzed to build predictors for process and states duration, which are the basis for creating alerts on ongoing processes.

In the following, the techniques developed for each of these goals are illustrated.

5.1 Variant analysis

Process mining tools, such as Apromore[2], allow the analysis of variants of a process and the identification of critical situations. Considering sequences of states, variants of a judicial process refer to the different paths that a case can follow during its evolution, considering its states as described in Section 3. Analysing these variants can help identify their recurring patterns, inefficiencies, and areas for improvement in the justice system.

As a first type of analysis, we focus on states and the transitions between states, ignoring internal events within states. By examining the sequences of states rather than individual events, the complexity arising from the multitude of events and their potential combinations is avoided.

Considering all defined processes, the average completion time of a process is 492 days, which is below the critical threshold of two years as defined by "Legge Pinto", an Italian law limiting the acceptable duration of a trial. From the analysis of the variants, 272 variants emerged, of which only 10 occurring more that 100 times during the period. Three main variants are identified, covering 66% of the cases, are discussed in the following, as they are the most representative of trials execution. On the other hand, all variants are considered as a baseline for the predictive techniques illustrated in the following, as they are associated to specific special cases.

[2] https://apromore.com/

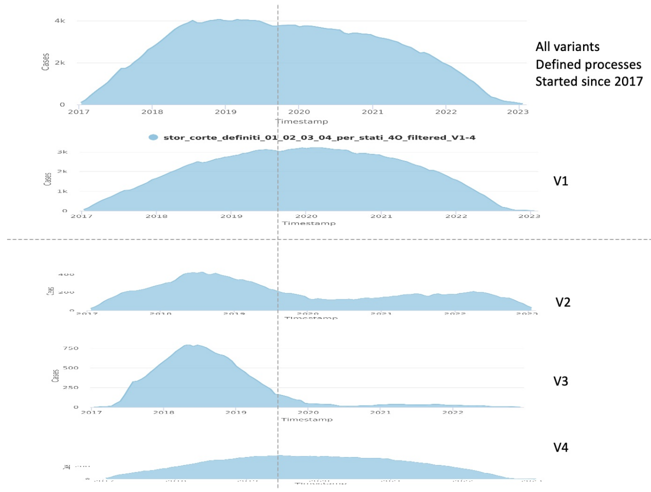

Fig. 5. Evolution in time of the main variants

- The most frequent variant (Variant 1 - 48% of the cases) reflects the expected process execution as illustrated in Section 3 and the average duration of trials in this case is 573 days.
- The second and third variants, with frequencies around 10% and 8% respectively, represent shortened paths, having respectively one and two states less than the main variant. These shortened routes reflect simplified legal procedures.

The rest of the variants, although less frequent, represent more particular cases that can offer useful information on specific situations or exceptions to the standard.

An interesting aspect from process mining is when the variants have been occurring in time. In Fig. 5, overall trends for variants are represented. First, considering all cases, the number of running cases over time is plotted. If the number of arriving cases is constant, this plot appears to be a Gaussian-like shape, as in the beginning we have only completed cases starting in the first year (previous years are not part of the dataset), and at the end fewer cases reach a completion stage. As it can be seen in the figure, there is a declining number of incoming cases over time, which shifts the peak of cases a bit on the left. This is confirmed by official statistical data of the Court, which has a reduction of the number of cases of civil trials in the last years.

Considering Variant V1, we see instead a stable situation, with a small trend of increasing cases towards more recent times, representing that the principal variant is being adopted more and more in the Court. Variant V2 (a simplified process) has been more applied in 2018 and 2019, and its use is rising again in recent times, which is confirmed by interviews of a trends towards increasing the number of simplified processes when possible. Variant V3, instead, shows a different pattern, and it appears to be not applied (or dismissed) in recent times. Variant V4 shows again a regular pattern of use during the period of interest.

5.2 Identification of critical events

Process mining with Apromore can be useful also to analyze events. In this case the starting and ending times of tasks are associated to the occurrence on the event and to its registration in the system, respectively. While it is difficult to identify sequence of patterns in this way, as the number of possible variants explodes considering all the possible events, the tool allows identifying critical events that have a duration much higher than other events, i.e., for which the recording has been performed with a significant delay. These events can be further analyzed, in order to identify patterns in which delayed registrations may occur. Further analysis can support the investigation of the consequences of a delayed registration. For instance, in Fig. 6 two scenarios for delayed events are outlined. Both diagrams refer to a delay of registration of events (x-axis) compared to the time of occurrence of the following event. The diagram on the left shows a case in which the following event occurs after the registration of the previous event, possibly with a further delay. In the diagram on the right, it is shown that the following event is always immediately after the registration of the previous event.

This allows two types of hypotheses of problems that can be investigated in the process: in the case on the right, the two events are recorded at the same time, when the second event occurs. This is usually a case in which events occur at different times but they are recorded in the system as a batch. Groups of three or four events of this type have been identified and may imply problems in understanding the state of a process in the real world for an operator, as the system does not reflect the occurrence of events within a reasonable time (usually one or two days). The case on the left can be a case of blocking event. The following event can occur only some time after the delayed registration, possibly causing a delay in the process.

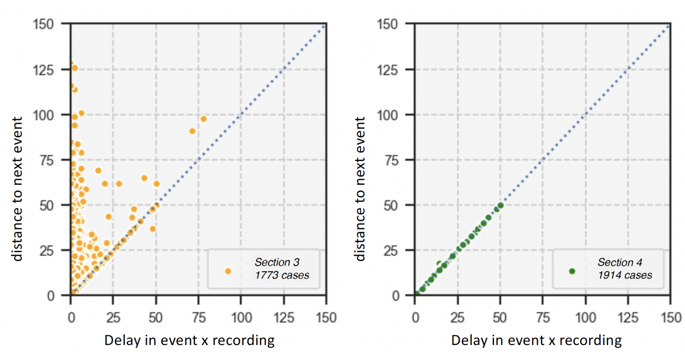

Fig. 6. Detecting batching events registrations

Another type of analysis of events is devoted to identify events that are critical for the process, as they always correspond to processes with a duration above the average. We focus here on delayed processes with legal implications, i.e., longer than two years. To achieve the goal of identifying what slows down the conclusion of a process we have used a number of different complementary approaches. These approaches were not carried out one after the other, but in parallel with the aim of not influencing each other. Once each of the analyzes was completed, we cross-referenced the results.

- We carried out an analysis of the types of possible events trying to find what could be the cause of a slowdown in the progress of the processes. Once an interesting subset had been identified, we checked whether they had actually led to a lengthening of times. The events thus identified were classified into different categories, in particular events may correspond either to simple postponements, or to requests for additional time, or requests for changes, or finally may be the cause of cycles in the workflow.
- We also carried out a purely numerical analysis, by considering only the times that characterize the processes and the events in them. What was done was to analyze the entire dataset by extracting the events that recur more frequently in longer processes and the events characterized by long waiting times before the next event.

The results of these two analyzes led to the identification of two sets of events that are potential causes of process delays; we realized that the intersection between the events extracted from these two approaches is extremely rich. As the two approaches were completely independent, their intersection was considered as particularly interesting and was subject to further investigation.

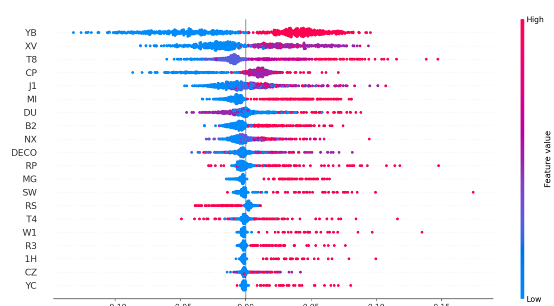

Fig. 7. Graphical representation of the Shapley values of a Random Forest model for the state "UT" Waiting for the first hearing (source: [8])

Fig. 7 shows preliminary results for a specific case; the events are sorted by importance and the coloured dots of the image are associated to each prediction of the legal processes. Blue dots represent the processes where the event is absent; violet and red dots (darker dots) represent the processes where the event is present once and more than once. The dots on the right side of the x-axis represent the predictions where the event influences positively the outcome; vice versa for the left side of the x-axis.

We can observe that in general red dots are on the right side of the x-axis, meaning that the presence of an event influences positively the total duration of the state; this was also expected a priori. Moreover, we can examine which events are more important by taking into consideration only the flexible ones, that can be optimized. Interesting events in the state of Waiting for the first hearing (UT) are "XV" and "MI". The first one represents the event of a lawyer who asks the permission to analyse the file of the process; the second one describes a postponement of the court hearing asked by one of the parties. Both events are important in terms of Shapley values and can be considered for optimisation. As a result, using machine learning models and techniques such as Shapley values and permutation importance, we gain a deeper understanding of how the events contribute to the duration of trials.

5.3 Predictive approaches for alerts

In the context of justice sector research, we propose an innovative tool for statistical prediction aimed at estimating the remaining duration of judicial procedures.

The proposed method takes into consideration several crucial parameters like the sequence of states traversed by a process (referred to as primitive), the matter and the object of the procedure, and the elapsed time from registration to the present. This information allows us to group similar processes, which makes it possible to calculate descriptive statistics and enhances forecast precision. The "primitive", an integral component of our approach, represents the sequence of states that a case has gone through, providing an overview of the path that a judicial proceeding has followed up to a given point. Once similar procedures have been identified, they are further aggregated based on the sequences followed from the last state of the primitive up to the ending state. For each group, the remaining mean time and standard deviation are calculated. This data serve as a foundation for the creation of an intuitive graphical representation offering remaining time estimates for each prospective path.

The opacity of the horizontal bars in the graph corresponds to the relative frequency of a particular continuation variant among similar cases, visually portraying the uncertainty linked to each prediction. In terms of estimating a procedure's remaining duration, we suggest two methods. The first is a frequency-weighted average of the mean remaining days for each type of continuation variant. Alternatively, the remaining time prediction could be the average remaining time of the most frequently occurring continuation variant.

Our statistical projection tool provides both a graphical and numerical depiction of the potential future paths of a proceeding. This representation empowers users to apply their expertise and experience to determine the most probable trajectory for a specific case and to refer to the relevant average remaining time, along with its confidence intervals.

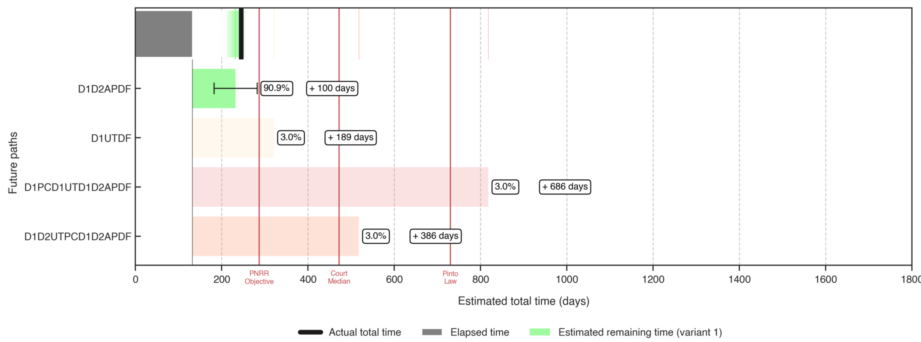

Fig. 8. An example illustrating a graph depicting the forecast remaining time for each variant of process continuation. Green bars indicate objectives are met. The intensity of red coloured bars indicates the severity of the delay.

Finally, this tool's incorporation into a management dashboard allows for real-time monitoring of case progression at the court of appeal. This feature is particularly crucial for setting alarms that trigger when predetermined time thresholds are surpassed, enabling proactive and timely oversight of procedural timelines.

To test this approach, we take the entire historical dataset and divide it into a training set and a test set. The training set consists of processes defined from 2017 until the end of 2021. The test set is composed of processes initiated after 2017 that are not yet completed as of January 1, 2022. Projections can be made when there is a state change in the process. On such occasions, the new state of the process becomes the last state in the sequence of the primitive. In some cases, the test set may contain processes for which there is no matching combination of matter, object, and primitive in the training set. In such cases, a rollback operation is performed, and the projections are calculated based on how the processes with that state continue after its occurrence.

The training set consists of 12,571 processes, while the test set consists of 3,461 processes. In terms of estimating a procedure's remaining duration, we propose two methods. The first method, which calculates predictions based on the frequency-weighted average of the mean remaining days for each type of continuation variant, yields an absolute error of 64.03 days. Furthermore, the confidence interval generated by this method encompasses the actual total duration of the process in 80.6% of cases. On the other hand, the second method, which utilizes the average remaining time of the most frequently occurring continuation variant, results in an absolute error of 65.81 days. The corresponding confidence interval contains the actual total duration of the process in 75.9% of cases. These findings suggest that both methods can be effective in estimating a procedure's remaining duration, with the first method exhibiting slightly better accuracy. It is also interesting to examine how this statistic varies depending on the starting state for the projection. We expect to observe an increasing success rate for starting states that occur on average in more advanced stages of the process (see Fig. 9).

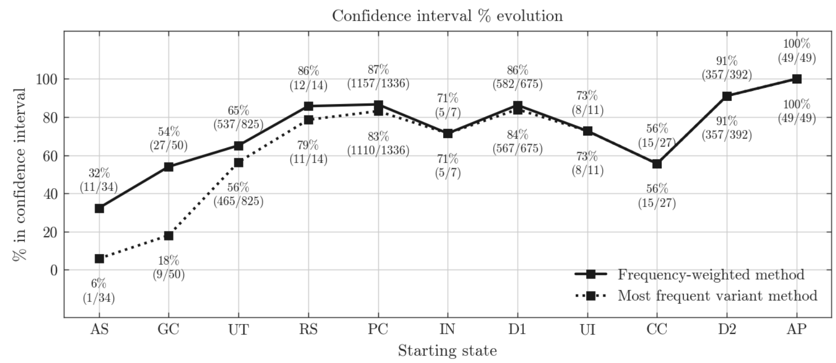

Fig. 9. Evolution of the percentage of cases whose actual total duration falls within the predicted confidence interval based on the starting state.

The graph in Fig. 9 confirms this hypothesis. The starting states for the projection are shown on the x-axis, while the y-axis represents the percentage of processes in each state for which the actual total time falls within the predicted confidence interval. The states are ordered on the x-axis by assigning a progressive index to the states within each process and sorting them based on the average of their indices. It is evident that giving an accurate prediction is more challenging for the early stages of the process (AS: "Attesa assegnazione a Sezione", GC: "Attesa designazione Giudice Rel./Collegio") due to the high variability in the future path. In such cases, it is more useful to visually evaluate the possible variants with their respective average duration as shown in Fig. 10 rather than condensing the information into a single value or confidence interval.

For states occurring later in the process, we can observe in Fig. 11 a gradual improvement in the success rate, with good values achieved for states which are included in most frequent variants, such as "Attesa udienza di Precisazione Conclusioni" (PC), "Attesa Deposito Conclusionali e Repliche" (D1, D2), and "Attesa deposito Provvedimenti" (AP). The same logic applies to the evolution of the absolute error concerning the states from which the projections are made. Once again, we observe a progressive improvement in predictions for stages that are generally more advanced within the process.

In conclusion, the results of our tests indicate that the ability of the tool to compute a confidence interval containing the actual total duration of the process is influenced by the starting state for the projection. However, we have noticed a progressive improvement in predictions as we move towards more advanced stages of the process. This trend can be attributed to the greater stability and predictability of activities at a more mature stage of the process. Therefore, we recommend carefully considering the starting point for projections to obtain

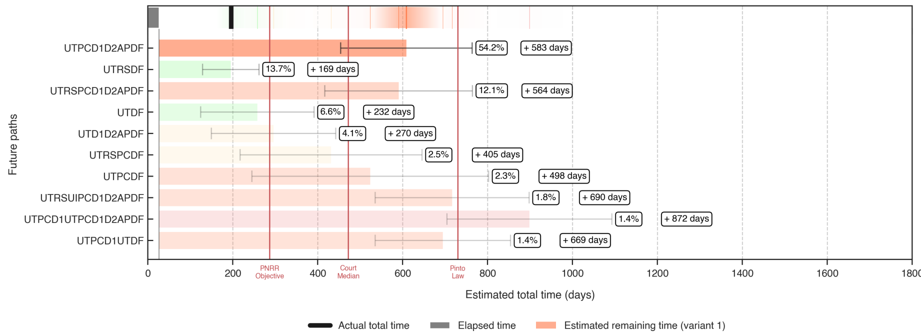

Fig. 10. Projections for a process duration in its current state of UT ("Attesa esito Udienza di Trattazione") with primitive "ASGC", meaning that the process has already gone through the awaiting of the assignment of the section (AS) and judge assignment (GC) phases. The graph demonstrates that making accurate predictions in the early stages is challenging due to the high variability of subsequent paths.

more accurate results. Additionally, the visual representation of possible variants and their average duration is particularly useful for the initial states of the process, where variability is higher. This approach allows a comprehensive assessment of potential process evolution, providing a more detailed and informative overview compared to a single prediction or confidence interval.

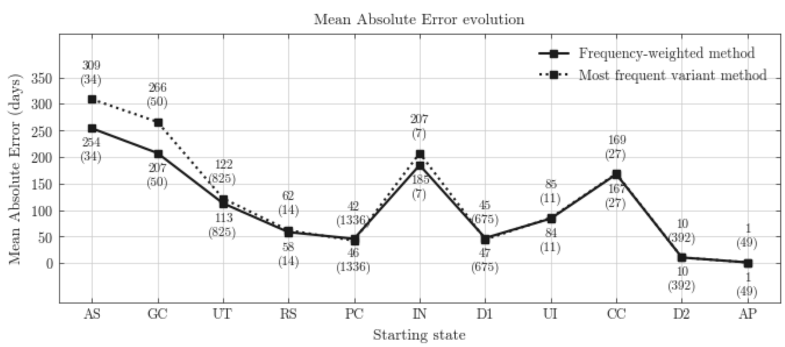

Fig. 11. Evolution of the Mean Absolute Error for each starting state.

6 Concluding remarks and future work

Our analysis approach provides a significant potential towards improved, informed judicial proceedings management. By intelligently utilizing historical data and employing advanced projection methodologies, we are equipped to provide dependable predictions that can aid judicial practices and steer strategic decisions. Nevertheless, it is important to acknowledge some inherent limitations and challenges. Primarily, our method operates on the assumption that future court proceedings will mirror past ones in terms of duration and paths. Nonetheless, justice is a dynamic system, susceptible to regulatory, organizational, and procedural changes that can significantly alter proceedings' timing and paths. Moreover, predicting the duration of proceedings may be influenced by other factors, such as the complexity of the case, individual judge decisions or involved parties' strategies. Data quality is another critical factor, as the accuracy of our predictions hinges heavily on the comprehensiveness, consistency, and correctness of the utilized historical data. Lastly, it is worth noting that predicting the duration of cases remains a probability-based estimate, and inherent uncertainty for each single case is always linked to the predictions. Therefore, predictions should be treated as guidance rather than an absolute forecast.

Acknowledgements This work has been supported the Italian Ministry of Justice PON project "Next Generation UPP: nuovi schemi collaborativi tra Università e uffici giudiziari. Per il miglioramento dell'efficienza e delle Prestazioni della giustizia nell'Italia del nord-ovest" and by the PNRR-PE-AI "FAIR" project funded by the NextGenerationEU program. The authors thank Ludovica Piro and Paolo Gerosa for their contributions in the first phases of the project.

References

1. van der Aalst, W.M.P., Schonenberg, M.H., Song, M.: Time prediction based on process mining. Inf. Syst. **36**(2), 450–475 (2011)
2. van der Aalst, W.M., Carmona, J.: Process Mining Handbook. Springer (2022)
3. Cugno, M., Giacomelli, S., Malgieri, L., Mocetti, S., Palumbo, G.: La giustizia civile in Italia: durata dei processi, produttività degli uffici e stabilità delle decisioni. Banca d'Italia (2022)
4. Dumas, M., et al.: AI-augmented business process management systems: a research manifesto. ACM Trans. on Management Information Systems **14**(1), 1–19 (2023)
5. Lashkevich, K., Milani, F., Chapela-Campa, D., Suvorau, I., Dumas, M.: Why am I waiting? Data-driven analysis of waiting times in business processes. In: CAiSE Conference. pp. 174–190. Springer (2023)
6. Martino, B.D., Cante, L.C., Esposito, A., Lupi, P., Orlando, M.: Temporal outlier analysis of online civil trial cases based on graph and process mining techniques. Int. J. Big Data Intell. **8**(1), 31–46 (2021)
7. de Oliveira, R.S., Reis Jr, A.S., Sperandio Nascimento, E.G.: Predicting the number of days in court cases using artificial intelligence. PloS one **17**(5), e0269008 (2022)
8. Pernici, B., Campi, A., Dilettis, M., Gerosa, P.: Why are Italian trials taking so long? a process mining approach. In: Proc. Workshop AI for Public Administration, Ital-IA, Pisa, Italy (May 2023)
9. Unger, A.J., Neto, J.F.d.S., Fantinato, M., Peres, S.M., Trecenti, J., Hirota, R.: Process mining-enabled jurimetrics: analysis of a Brazilian court's judicial performance in the business law processing. In: Proc. 18th Intl. Conf. on Artificial Intelligence and Law. pp. 240–244 (2021)