Ezio Ornato

An Experiment in Dating Documents through the Analysis of Watermarks: The Letter ‘P’ in Incunabula of the Low Countries

Summary: The systematic application of dating methods through the analysis of watermarks in a large number of documents not only requires the automation of comparison procedures (i.e. the matching of identical or almost identical watermarks), but also necessitates the rigorous and effective processing of evaluation criteria, the validity of which can be confirmed through statistical analysis. With this objective in mind, the present contribution sets forth the very encouraging results of an experiment conducted on examples of the letter ‘P’ collected in the *WILC* database (*Watermarks in Incunabula printed in the Low Countries*), which should facilitate the dating of a large percentage of early printed editions within a margin of error of less than t ± 12 months.

1 Problems associated with dating documents through watermarks

Leaving aside the physico-chemical procedures currently employed in the field of archaeology, the dating of documents through the analysis of watermarks is perhaps the only method based on objective criteria that can be applied in the sphere of the human sciences, despite being susceptible to an inevitable degree of inaccuracy. Although its application is necessarily limited to paper-based documents produced in the West in the Late Middle Ages, the method appears to hold great potential, given on the one hand the tens of thousands of undated manuscripts and printed books conserved in libraries, and on the other the remarkable degree of accuracy that seems, in theory, to be achievable.

Translated from the Italian into English by Mark Livesey. Original published as Ornato, Ezio (2004), ‘Un esperimento di datazione tramite le filigrane : Le “lettere P” negli incunaboli dei Paesi Bassi’, in Graziaplena, Rosella, with the assistance of Mark Livesey (eds), *Paper as a Medium of Cultural Heritage: Archaeology and Conservation*. 26th Congress IPH, Roma: Istituto centrale per la patologia del libro (Addenda, 5), 225–242.

https://doi.org/10.1515/9783110743838-007

Yet, despite their undeniable potential, the methods used for dating documents through the watermarks they contain have never been applied on a large scale. Furthermore, even when such methods are used in individual cases, the datings obtained do not have any independent value—indeed, their currency is always subordinate to the results obtained through philological and palaeographical studies, so even at the best of times they only constitute supplementary evidence. It is scarcely necessary, then, to underscore the fundamentally unscientific and risk-laden nature of such methods: the concordance of doubtful conclusions does not, *per se*, increase their accuracy, since the margin of error is not quantified. Several reasons behind lie behind this paradoxical state of affairs, which are as follows:

– Comparison procedures are onerous and have to be carried out by hand on multiple catalogues.
– The chances of finding watermarks that are identical to the one being dated are slim.
– Working methods are of low accuracy and have not been significantly improved on since Charles-Moïse Briquet first presented the fundamentals of his approach in the early twentieth century.
– The potential for making serious errors exists, due to the impossibility of establishing reliability criteria that are statistically quantifiable.

Is it possible, then, to proceed any further? Happily, the answer to this question is in the affirmative.

In 1996, at the IPH congress held in Leipzig, the *Progetto Carta* collective[1] proposed a comparison and dating method based on morphometric criteria—i.e. type and size of watermarks—applied to a corpus of 'scales enclosed within a circle' specimens surveyed in the inventories of Briquet and Piccard, but above all in a corpus of a few dozen Venetian incunabula studied during the course of *Progetto Carta* itself.[2] If the results obtained from the inventories had already proved quite satisfactory, the datings of printed books delivered highly accurate results—to within ± 1 year in the best cases—with a margin of error of less than 5%.

The conclusions reached through the experiment conducted on the 'scales enclosed within a circle' probably appeared a little too optimistic to some scholars, chiefly because of the chronological distribution of incunabula included in the corpus, which was heavily concentrated on the years 1493 to 1494, and

1 Concerning this research project, which was instigated by the ex-director of the *Istituto Centrale per la Patologia del Libro* in Rome, see Ornato et al. 2001.

2 Ornato et al. 2000, 61–75.

therefore tended to minimise the variation between estimated and actual numbers. The new experiment—which no longer focused on the 'scales' watermark, but instead on the letter 'P' in a corpus of incunabula widely distributed over the last three decades of the 15th century—confirmed, and even improved on the results obtained in the previous experiment.

The experiment was only made possible thanks to the existence of a new database—*WILC (Watermarks in Incunabula printed in the Low Countries)*—which now contains more than 4,500 watermarks harvested from incunabula of the Low Countries (in the broad sense, i.e. Belgium and Holland). The databank was developed by Gerard Van Thienen, and supported by the Koninklijke Bibliotetheek in The Hague. Reproductions of the watermarks, most of which are radiographs, have been made freely accessible to scholars on the library's website.[3] The databank presently contains more than 1,700 letter 'P' watermarks.

Apart from the strictly watermark-related aspects of the database, it should be emphasised that *WILC* represents a very rich virtual source of knowledge and information, just as much on Western paper as on the 'primordial' days of the printed book. A visit to the website, where the purpose and characteristics of the database are clearly explained, as well as, of course, instructions for its use, will certainly be more worthwhile to the reader than any written description. Without getting lost in minutiae (which in any case are often pointless, given the existence of images), suffice to say that *WILC* adopts an approach precisely opposite to that of Piccard's catalogue, which is very rich in images but equally poor in providing information on the characteristics of watermarked leaves. In fact, the Dutch website provides a wealth of indispensable information for the benefit of scholars,[4] added to which is all the relevant data concerning the particular incunabulum that the leaf originates from.[5] When an exact date is not provided, a rough estimate is made, together with a justification for it.

But there is more: Gerard Van Thienen had already embarked on the monumental and painstaking task of comparing a large number of watermarks by hand, work which led him to establish some 'equivalents categories' (about

3 https://watermark.kb.nl/page/index/title/Introduction (last access 07/09/2021).

4 Identification of the imprint side (i.e. the side of the sheet that was in contact with the woven surface of the mould during the manufacture of the sheet), which makes it possible to determine the position of the watermark on either the right- or left-hand side of the mould; the orientation of the motif in relation to the vertical axis in the presence of asymmetry (the 'belly' or 'bowl' of the letter 'P') on the right or on the left; the gap between two chain lines where the watermark is located; the distance between the chain lines that surround the watermark; and finally, the height of the watermark itself.

5 Format, place and date of publication, information about the printers.

which more will be said later on) which group together watermarks that appear to be identical, or are at least very similar. This important preliminary contribution made it possible to provide considerably more detailed verifications than had been possible in the experiment on the 'scales enclosed within a circle' motif.

All methodological verification experiments on the dating of documents through the watermarks they contain are based on a simulation procedure: the assumption is made that the date of a given sheet containing a watermark is unknown, whilst in reality it is, in fact, known, thus making it possible to confirm, *a posteriori*, the accuracy of the estimated date.

In broad terms, the following steps are required in order to carry out the simulation procedure:

– The establishment of a 'reference group', which is to say a corpus of watermarks of sufficient size so as to be significant.
– The definition of a gauging criterion to establish the similarity—or more easily, dissimilarity—between two watermarks, to be represented by a quantitative index.
– The application of the gauging criterion to all the watermarks contained in the corpus. In practice, this involves comparing each watermark in the corpus with all the others. If *N* is the size of the corpus, a total of *N* x (N-1)/2 comparisons have to be carried out.
– For each comparison, two pieces of data are obtained, namely an index of dissimilarity (or 'morphometric gap'), and a 'chronological gap', which is to say the difference in absolute terms between the dates of the two documents containing the watermark.
– The statistical distribution of the chronological gaps in relation to the morphometric gaps will show the probability that, for a morphometric gap 'MG', the chronological gap 'CG' will be either higher or lower than a predetermined value.
– The assumption is made that the results obtained through the simulation carried out on watermarks whose dates are known (i.e. in dated books and documents) are equally valid for watermarks whose dates are unknown. In other words, for a given 'MG' value, it should be possible to deduce the likelihood that the unknown date of a watermark lies within the time range arrived at experimentally, for that value, in the reference group whose date is known.

The above being said, both the dating procedure and the simulation experiment have to be carried on the basis of some specific criteria. The fundamental postulate that guarantees the validity of the method is the fact that two sheets of paper originating from one and the same mould, but which appear in different written documents, cannot, statistically speaking, be very distant from each other in chronological terms. In fact, the average 'CG' among the various attestations is dependent on the duration of the mould's 'usage cycle' (i.e. the length of time it was in service), and the duration of the 'consumption cycle' of the paper produced from it (i.e. the timeframe within which all the sheets made with that particular mould were consumed).

These two parameters—which can vary over time, and also in synchrony according to different rates of consumption—jointly determine the maximum theoretical degree of accuracy that the method can achieve when one is estimating the date of a sheet devoid of chronological data. However, since a document to be dated is not usually composed of a single sheet of paper but instead consists of an entire volume, the level of accuracy achieved by examining a single sheet can be increased when one and the same volume is made from sheets that emanated from different moulds. The concomitant presence of different sheets within one and the same stock of paper necessarily lasts a shorter length of time than the consumption cycle of each of the individual sheets.[6] In particular, it should be emphasised that a document's date cannot be set earlier than the moment in time when the most recent mould (represented by the sheets emanating from it) entered the production cycle.

A second postulate—whose significance has already been confirmed during the experiment performed on the 'scales' motif—presupposes that, even in the absence of a perfect match, the 'similarity' between two watermarks is, within certain limits, negatively correlated to the average chronological gap between their attestations in different volumes: in other words, the more two watermarks resemble each other, the more their attestations tend to be chronologically close to each other.[7] This second postulate makes it possible to justify the use, in the dating procedures, of non-identical watermarks, and also makes it unnecessary to worry about knowing in advance whether or not two watermarks are identical.[8]

6 The principle had already been well understood and applied by Briquet; see Briquet 1968 (1907).

7 Needless to say, this does not imply the opposite, which is to say the more different the watermarks are, the more chronologically distant from each other they will be.

8 It should be borne in mind that in cases where similar watermarks were used contemporaneously on different moulds employed in the same paper mill, the accuracy of the estimate will be almost the same as that in a situation where the watermarks are identical. If, on the other hand, one is dealing with watermarks sewn on to moulds scattered over the same time vector (i.e. new moulds which substituted their worn out predecessors), the level of accuracy will clearly be lower. The problem here is that one is not normally able to differentiate between the two

If one were able to have at his or her disposal a ‘perfect’ database containing all the attestations of each and every occasion a particular mould was utilised over a period of time, the level of accuracy achieved by the dating method based on watermarks would in every instance match the theoretical maximum degree of accuracy. For a single watermarked sheet, this would depend on the length of time the usage and consumption cycles lasted. For an entire volume, it would also depend on the number of different sheets present in the document to be allocated with a date.[9]

Alas, the aforementioned ‘perfect’ database does not exist. This is principally because such a database simply *cannot* exist, given the vast number of sheets of paper that have been lost over the course of time. However, if one considers the methods currently employed in the field of statistics, it goes without saying that it would not in fact be necessary to have at one’s disposal all the attestations which have ever existed. Instead, for each mould, a sufficiently representative number, which is to say the number that would be necessary in order to allow one to reconstruct (in an accurate way) the chronological distribution curve in question, would suffice.

In all likelihood, even in the best of cases this criterion could never be fully satisfied—indeed, in order to obtain a reliable curve, it would be necessary to have at one’s disposal at least thirty independent attestations for each mould. However, if one considers that the total production emanating from a single mould approached a million sheets, it is quite likely that at least one dated attestation of every mould used in the paper mills still exists today and, not infrequently, a relatively large number of attestations exist. One has only to consider the fact that Theo Gerardy did not have too much trouble in identifying all the alternating moulds that were used over a period of more than a hundred years at the Arensberg paper mill.[10] On the other hand, the comparisons carried out by Gerard Van Thienen demonstrate that one and the same watermark can appear at least ten times in different editions, and to these attestations can be added those originating from archives. All this leads one to deduce that for each ‘uncharacterised’ sheet (i.e. a sheet not endowed with a date and geographical

situations, except for when one is dealing with obviously ‘twin’ watermarks (a situation, moreover, which is not always easy to identify with the utmost certitude).

9 For a given watermark, the chronological distribution curve for attestations will be strikingly asymmetrical. It will rapidly reach a maximum when the mould ceases to be serviceable, and therefore to supply the market. Subsequently, the curve shows a gradual diminution and heads towards zero little by little as the stocks in circulation gradually dwindle.

10 See Gerardy 1964.

origin), somewhere or other at least one corresponding characterised attestation must exist, which should, in theory, make at least one verification possible.

It is fairly obvious that the best approximation (in concrete terms) of the original situation, which is to say all the individual attestations of use that are currently known, would in itself be sufficient to produce datings within a time range very close to the maximum accuracy permitted by the method. Needless to say, we are very far from enjoying this ideal situation. The 'watermark universe' which is currently available to us in the form of reproductions—the bulk of which, unfortunately, consist in tracings—is still inadequate, and indeed would remain so even if one succeeded in combining in a single database all the material surveyed up to the present, both published and unpublished.

If the accuracy of dating depends on the objective factors listed above, which are bound by insuperable limits and are entirely independent from the wishes of scholars, it also depends on the richness of the dated reference group of comparison and its geographical and/or chronological relevance with respect to the material to be dated. Unlike the former, these latter factors are dependent on human initiative. If the gridwork in the reference group points to consumption areas which lie outside those relevant to the document to be dated, the number of 'effective' comparisons (i.e. those which correspond to morphometric gaps that are sufficiently small) will rapidly diminish, whereas if the gridwork is too wide, their quality (i.e. the degree of similarity revealed between the watermark to be dated and those present in the reference group) will decrease.

Consequently, even when the material to be dated corresponds perfectly, on the chronological level, with the reference group, the accuracy and reliability of a dating can never exceed the limits set its richness, and therefore what we can define as the 'degree of internal datability'. This 'degree of internal datability' can be ascertained through experimental tests; indeed, the results of such tests are used to define the efficacy, reliability and accuracy of the dating procedure.

How, then, does the 'degree of internal datability' test work? We have already seen how a comparison with all the watermarks in the reference group makes it possible to determine the statistical distribution of the chronological gaps 'CG', corresponding to the different values of the morphometric gaps 'MG'. The said distribution can be represented by a cumulative frequency curve, which at the same time will define the accuracy of the dating and its reliability.

For example, let us suppose—needless to say, we are dealing with fictional values—that the 'MG' category is 0–5, and that 90% of the 'CG' falls within the ±6 months range, while 95% falls within the ±12 months range. This means that if the comparison between an undated watermark and one or more watermarks in the reference group produces an 'MG' that ranges from 0–5, one can state that its probable date is the same as the comparison watermark ±12 months, with a bilateral margin of error of 10%. However, if the 'MG' value is higher, in the range of 6–10, the experimental test will necessarily produce less accurate results, e.g. ±18 months, with a margin of error of 5%. Finally, if the 'MG' value is even higher—e.g. ranging from 20–30—the level of accuracy will fall to ±84 months, and therefore the comparison will be practically useless for the purposes of dating.[11]

The advantage of this empirico-statistical procedure is that on the one hand it provides the possibility of automatically deciding whether the results of the comparison between two watermarks are usable for dating purposes, and on the other that of quantifying—and adjusting according to one's needs and specific concerns—the accuracy of dating and margins of error. Once the characteristics (in due course we shall see which) of a watermark are fed into the computer and automatically compared with all the others in the reference group, and the results then evaluated against the dated material within that group, the researcher has nothing further to do, other than, if necessary, interact with the options proposed by the software.

2 An experiment in dating on the '*WILC*-Letter P' corpus

The journey towards the 'promised land' is, however, full of pitfalls. The first problem to tackle consists in obtaining, with as little wasted energy as possible and the maximum degree of accuracy, an 'MG' measurement that has the necessary characteristics. Given that no procedure currently exists (despite ongoing research into the same) which makes it possible to automatically define the morphological and dimensional properties of two drawings, whilst at the same time providing a numerical assessment of their degree of similarity, one is constrained to work in a purely manual way. Such an approach is hampered by severe constraints.

A necessary condition for two watermarks to be judged identical is that they be exactly superimposable upon each other. Using software that enables one to manage different 'levels'[12] makes it possible to manually confirm that two motifs can be precisely superimposed upon each other. However, the manual

11 This degree of accuracy is achievable using other methods, above all when one is dealing with printed editions.

12 The 'levels' can be defined as two or more 'strata' of the image rendered 'transparent' by the software. The 'levels' can then be superimposed on each other so as to form a single image.

procedure—which is useful for confirming the results of research aimed at grouping together different attestations of an identical watermark—is a long and drawn-out process when it has to be systematically applied to hundreds of different watermark designs. On the other hand, it is impossible to easily, objectively and rapidly measure the morphometric gap 'MG' when two images cannot be exactly superimposed upon each other.

The most practical solution consists in avoiding a graphics-based approach, and instead employing a numerical and dimensional representation of the watermark in question, consisting of a series of dimensional measurements. The comparison between two numerical representations will provide, *ipso facto*, an immediately usable morphometric gap. For a given watermark motif—such as the letter 'P'—in the preliminary phase a certain number of key points are pre-selected whose reciprocal positions are then measured in all the watermarks to be compared. Needless to say, it is necessary for the key points to be fairly numerous in number so as to minimise the risk of chance matches, but at the same time there should not be too many of them in order to avoid hampering the measuring procedure.

In theory, the ideal approach would be to compare the coordinates of one and same point in relation to a fixed reference point in two variants of a particular watermark design.[13] The weakness in this type of approach lies in its vulnerability to two phenomena that often affect watermarks, namely their rotation around a geometric barycentre and their 'drift' towards one of the two adjacent chain lines. An alternative procedure, which is simpler and quicker to carry out, consists in measuring a certain number of arbitrarily pre-selected spaces between pairs of points—always the same ones, of course. However, the latter method is less selective than the former.[14] In order to measure the gaps between each pair of points—10, for example—it would be necessary to make 45 measurements, which besides (obviously) being impractical, would of course result in a large surfeit of data.

The solution settled on consisted in pre-selecting 13 points called 'nodes'—located at the meeting point of two segments of the letter 'P'—and then measuring 10 distances, with all 13 of the points included at least once. Given that radiographs and betagraphs are susceptible to a certain number of 'visual impurities'

13 The coordinates of a point are defined by its distance in relation to a system of Cartesian axes, or by the gap in relation to the reference point and the angle formed by the segment which unites the two points with one of the two Cartesian axes.

14 In fact, when presented with four points—A and B, and A' and B'—arranged in the same way in two different watermarks, if the two points A and A' coincide, the distances A-B and A'-B' can clearly be the same even if points B and B' do not coincide. One has only to think of an isosceles triangle, in which points B and C in its base, despite not coinciding, are positioned at equal distances from the triangle's apex.

which sometimes make it impossible, when taking measurements, to define the aforementioned points with sufficient accuracy (such as in the case of the meeting point between the two 'legs' of the 'Gothic P'), a certain number of nodal points were eliminated at the outset. Here, it should be emphasised that these nodal points probably do not represent a purely abstract entity: the rather sketchy information we possess on watermark technology—all very recently acquired, besides—describes pegs driven into a wooden board, around which a wire strand was arranged so as to form the watermark's motif.[15] This observation is important, since the watermark makers could, if necessary, save a peg system and reuse it at a later date. This phenomenon represents an important 'disruptive' factor.

Irrespective of this possibility, it should not be forgotten that the representation of a design using only 13 geometric points constitutes a drastically reductive step which is highly vulnerable to the so-called 'first type' error, namely considering two motifs to be identical when in fact they are not. Conversely, errors in measurements caused by insufficient definition of the nodal points raises the risk of the 'second type' error occurring, namely considering two watermarks to be different when in fact they are identical. The first type of error introduces a 'noise' capable of seriously skewing the results of the dating procedure, while the second eliminates useful information and could potentially hinder the dating procedure or reduce its degree of accuracy.

In addition, the selection of 13 nodal points and 10 gap measurements made it necessary to choose from a range of preliminary options. The most important among these concerns the inclusion or exclusion of 'superfluous information', which is to say of all the details of the design that are of no great consequence to the basic structure of the letter 'P'. The exclusion of 'superfluous' details increases the risk of the 'first type' error occurring, since two watermarks that are characterised by the same essential structure, but which can be distinguished from one another at first glance due to the presence or absence of 'superfluous' elements, will be judged identical by the computer. On the other hand, inclusion of such details does not come without certain drawbacks on the historical level, since it eliminates *per se* the possibility of identifying the presence of basic construction patterns, and as a consequence the possibility of recognising the re-use and/or sustained use over time, or in a particular geographical setting, of the same watermark design. Since it is important to reduce to a minimum the risk of the 'first type' error, one has by necessity to align a metrical representation with a somewhat simplified typological classification system, which can then be usefully applied as a discriminating factor.

---

15 See Grosse Stoltenberg 1965, 73–79.

From the aforesaid, it becomes clear that two designs, judged as identical from a morphological point of view, can in fact be related to two different material objects based on the same archetypal model, but sewn on to different moulds and separated by a considerable period of time. For this reason, the typological classification has to take into account certain factors, namely the position of the watermark on the mould (on the left-hand or right-hand side), and the orientation of the motif (normal or specular).[16] In a scenario of this kind, another discriminating parameter can be found in the density of the laid lines, which makes it possible to reveal even minor differences in a mould's weave pattern, and therefore also to eliminate the misleading effects of any chronological 'mirages'. In the matter in question, the assessment of laid line density was greatly aided by an automatic counting program devised by Vlad Atanasiu.[17]

The metrological distance between two watermarks is obtained in a very simple way: the difference between the pre-selected distances in the two watermarks placed in comparison is automatically calculated by the computer. The difference between the figures is then squared (so as to prevent variations in the opposite direction from being cancelled out). The metrological distance is equal to the sum of the variations calculated in 10 segments. Just how reliable, then, are such measurements?

First, the problem of errors occurring during the measuring procedure has to be considered. In this regard, the working conditions are almost perfect thanks to the use of graphics software such as Photoshop. Such software enables one to work on images at life-size and to measure, on the computer screen, the distance between two points to within 0.10 mm. In addition, by using the Photoshop option called 'levels', it becomes possible to superimpose any two images upon each other with a very high degree of precision for the purposes of verification.

One naturally assumes that the measurements obtained from radiographic reproductions should be more reliable than those obtained from motifs of the watermark designs found in the catalogues. In reality the situation is not, in fact, quite so rosy, for the following reasons:

– A watermark's motif can be affected by various inaccuracies. Indeed, poorly defined areas are often observed in proximity to some of the nodal points, and these inevitably result in discrepancies in measurements.

16 All these elements are noted in the *WILC* database, but unfortunately are absent from the two extensive watermark catalogues that have been published to date.

17 Atanasiu 2004, 172–184.

- The radiographic image may not be perfect, due to the presence of decorated initials situated over the watermark motif, or simply because the image is out of focus.[18]
- The limited skill or dexterity of the operator can cause problems. However, the impact of this issue can be assessed, later on, by measuring the same image a second time.[19] Tests have been carried out which demonstrate that the variation between two successive measurements on the same image are entirely within the bounds of acceptability, on condition that the radiograph in question is of optimal quality.

All of the above drawbacks result in a sort of 'background noise' which cannot be eliminated. Consequently, the morphometric gap between two watermarks can never be valued at zero. One may imagine that the schematic drawings provided in the catalogues would make it possible to eliminate, or at least reduce, errors in measurements, but it should not be forgotten that one is often presented with simplified sketches (obtained through tracing) of the original watermark, and therefore errors start occurring at the outset, and worse still, cannot be identified.

As has already been mentioned earlier on, the *WILC* database currently contains roughly 1,700 letter 'P' watermarks. The ultimate goal of the project is to date *all* the editions published in Belgium and Holland whose dates are unknown, making use also of other watermark motifs that are well represented in the *WILC* database. However, since the experiment was focused above all on verifying the comparison method's potential, it was conducted under the most favourable conditions. Therefore, only 625 letter 'P' watermarks reviewed in 88 folio editions explicitly dated by their printers, and belonging to 'equivalents groups' previously identified by Gerard Van Thienen,[20] were measured. For the 625 chosen watermarks it was necessary to carry out 6,250 manual measurements and 195,000 automatic comparisons.

What, then, on the statistical plane, is the theoretical dating potential of the database '*WILC*-letter P'?[21] In order to answer this question, attestations of each

18 The lines forming an image can become 'swollen' when an image is out of focus, which results in imprecise measurements being taken.

19 Although not the same watermark in two different sheets, because the use of the mould, deformations in the watermark, and the handiwork of the vat workers can all exert an influence on a watermark's characteristics.

20 'Identical' watermarks found in one and the same edition were not included because, needless to say, their simultaneous presence is of no consequence.

21 In other words, the degree of accuracy it is possible to achieve for individual watermarks when the comparison is limited and includes only specimens that are considered 'identical'.

watermark within an 'equivalents group' were arranged in chronological order, and for each group the chronological gaps between the oldest and most recent attestations were measured.

Tab. 1 presents the results of the operation: in essence, the chronological gap within one and the same group exceeds 36 months in only 5% of cases. This means that when it is possible to compare an undated watermark with at least two others belonging to the same 'equivalents group', the most probable date will be provided by the arithmetic mean lying between the two temporal extremes, with a level of accuracy of ±18 months. It goes without saying that this result cannot be interpreted as an absolute fact—indeed, this level of accuracy relates specifically to the rate of paper consumption on the part of printers in the last thirty years of the 15th century in the Low Countries.

Probability interval between the earliest and latest examples of two 'equivalent' watermarks (tab. 1)

Months range	Cumulative percentage of watermarks
0–6	31.56%
6–12	55.67%
12–18	74.82%
18–24	86.52%
24–30	91.13%
30–36	94.33%
36–42	96.10%
42–48	96.45%
48–54	97.52%
54–60	97.87%
60–66	98.94%
66–70	99.29%

Tab. 1: Probability interval between the earliest and latest examples of two 'equivalent' watermarks

2.1 Description of the morphometric method and its efficacy

Alternatively, let us suppose that in the absence of any kind of preliminary comparison, nothing whatsoever is known about the typology and morphology of the watermarks, and therefore one has to proceed in a different way, with the following objectives in mind:

– The morphometric comparison procedure must be rapid, thorough and systematic, which means it has to be automated.
– For each watermark, the probable dating has to be suggested by the computer and interactively accepted by the scholar/researcher.
– The dates adjudged to be most likely for each simultaneously present watermark in a given edition have to contribute to its dating, thanks to the application of an automated and interactive algorithm.

As part of the experiment conducted on the letter 'P', and thanks to the descriptions provided in the *WILC* database and an autopsy of the images, all the watermarks were initially classified on the typological level. The classification was based not only on the characteristics of additional elements or the presence/absence of the same, but also according to the position and orientation of the watermark motif in relation to the mould. Information on the density of the laid lines was added to the typological data. The ten measurements were then fed into a database, together with information about the incunabulum in which the watermarks were found, with particular regard to its date. Next, the measurements obtained for each watermark were grouped into hierarchical classes and arranged in descending order (i.e. from the best to worst result) so as to ease data processing.

In a preliminary phase, all the watermarks which resulted in at least one comparison showing a morphometric gap of <21 were deemed 'useful watermarks' for dating purposes. The results obtained are presented in Tab. 2. It should be emphasised that the <21 morphometric gap criterion (MG<21) alone proved to be entirely worthless. On the other hand, the concordance of several criteria—which correspond to a situation of presumed morphometric similarity (the last column in the table)—makes it possible to improve the estimate, raising it to a ±30 months level of accuracy, with a margin of error of 5%. However, the watermarks that simultaneously fulfil all the necessary criteria represent only 28% of the total.

Results of morphometric comparisons made between 'useful watermarks' (for the purpose of dating) (tab. 2)

Morphological distance	indifferent	0	0	0	0
Morphometric distance	<21	<21	0–5	0–5	0–5
Chain lines distance	indifferent	indifferent	indifferent	≤ 0.2mm	≤ 0.2mm
Watermark position	indifferent	indifferent	indifferent	indifferent	identical
Orientation of 'P'	indifferent	indifferent	indifferent	indifferent	identical
% of useful watermarks	76.4%	51.6%	45.3%	33.2%	28.3%
Months error +/- (prob. 5%)	114	90	70	42	30

Tab. 2: Results of morphometric comparisons made between 'useful watermarks' (for the purpose of dating)

A closer analysis demonstrates that the criteria which exert the greatest influence on the accuracy of dating are the morphometric gap 'MG' and the density of laid lines. In cases where no more than two exceptions are permitted among the criteria, the level of accuracy remains practically unchanged, whilst the percentage of watermarks considered useful for dating purposes rises to 58%.

The ±30 months (i.e. ±2.5 years) level of accuracy cannot be considered very satisfactory. This relatively unimpressive result is chiefly due to an insufficient differentiation of the typological elements and the low degree of variability in the design of the letter 'P'.[22] It should not be forgotten, however, that the objective is to date not just a single sheet but an entire printed volume which almost always contains watermarks of different types (i.e. motifs) and variants within each motif type. Tab. 3 presents the results obtained from 78 editions, representing about 90% of the books included in the experiment.[23]

22 If this second factor is partly intrinsic in nature—a letter 'P' does not hold much potential for creative interpretations—it also suggests a prolonged use of the same construction patterns. In other words, for the letter 'P' there does not appear to exist a deliberate tendency to distinguish watermarks of same the design in an ever more elaborate way, whilst the opposite is observed in the watermarks found in papers originating from the Garda region used by Venetian printers.

23 The remaining 10% did not contain any letter 'P', or at least none which were useful for the purposes of dating.

Edition datings. Degree of accuracy depending on the estimation method employed (tab. 3)

Months error	Truncated averages	Overall averages
-30 -24	1.28%	2.50%
-24 -18	3.85%	5.13%
-18 -12	5.13%	7.69%
-12 -6	2.56%	7.69%
-6 -0	24.36%	23.08%
0 +6	38.46%	26.92%
6 +12	15.38%	17.95%
12 +18	6.41%	3.85%
18 +24	0.00%	1.28%
24 +30	2.56%	2.56%
36 +42	0.00%	1.28%
Total	-18 / +18 = 92.31%	87.18%

Tab. 3: Edition datings. Degree of accuracy depending on the estimation method employed

The final column in the table presents the results obtained when the likely date of an edition is reckoned to be the arithmetic mean of the dates of all the letter 'P' present on its leaves. The confidence interval at approximately 95% is ±24 months. This result, although not entirely negligible, is not altogether satisfactory. The situation improves if the mean is calculated by subtracting, respectively, the earliest and most recent dates of the simultaneously present watermarks (i.e. a truncated mean), just like in ice skating and ski jumping competitions. In the case of watermarks, this expedient makes it possible to mitigate the effect of what we can term 'motif inertia', which is to say the 'recycling' of a motif over an extended period of time. The application of this expedient makes it possible to reduce the dating range to around ±18 months (the shaded area in the table).

A further analysis of the data discloses the fact that dating errors are closely linked to the number of usable watermarks simultaneously present in a given edition. In particular, the mean error decreases considerably as soon as three or more different watermarks are available. The application of this criterion (Tab. 4) allows for the dating of 50% of the volumes to within -6/+12 months, or in practical terms, ±9 months.

Dating of editions containing more than three useful watermarks (tab. 4)

Months error	Actual percentage	Editions
-6 -0	34.09%	15
0 +6	40.91%	18
+6 +12	20.45%	9
+12 +18	4.55%	2
Total	-6 / +12 = 95.45%	44 = 50%

Tab. 4: Dating of editions containing more than three useful watermarks

This new result is truly excellent, and can only be improved on—not because the method implicitly allows for greater accuracy, but rather because it makes it possible to significantly increase the percentage of datable editions. Indeed, on the one hand, in the case of the letter 'P', the watermarks dated through the application of this method could be used, with due caution, to date other editions, whilst on the other hand other watermark motifs sufficiently represented in the *WILC* database can be subjected to the same experimental investigation.

On a trial basis, a simpler procedure was applied to the same 625 watermarks, taking as an established fact the similarities between various elements within one and the same 'equivalents group' observed during the manual comparisons carried out by Gerard Van Thienen. Excluding, as before, the dates at the extreme ends of the temporal scale, all the watermarks of a single group had allocated to them the average attestation date of the group itself, and the date of an edition was then estimated by calculating the truncated mean of all the watermarks present in it.

Tab. 5 shows that 90% of the editions can be dated with a 95% degree of likelihood to a time range of -12/+18 months (shaded area), or in practical terms, ±15 months. The result is better than that obtained using the morphometric method. There are two reasons for this: (a) the greater reduction in the 'design inertia' effect, and (b) the almost certain similarity among the watermarks belonging to one and the same group. It should be borne in mind, however, that the manual work required to sort 'similar' watermarks into groups is very time-consuming.

Dating of editions through 'equivalence groups' (tab. 5)

Months error	Actual percentage	Editions
-36 -30	1.28%	1
-18 -12	3.85%	3
-12 -6	8.97%	7
-6 -0	26.92%	21
0 +6	34.62%	27
6–12	15.38%	12
12–18	8.97%	7
Total	-12 / +18 = 94.90%	78 = 88.60%

Tab. 5: Dating of editions through 'equivalence groups'

If one considers only the editions that include three different equivalents groups (Tab. 6), the result obtained will improve considerably, with the range decreasing to ±8 months, with scarcely any margin of error.

On the other side of the coin, since the simultaneous presence of numerous groups is even more improbable than that of numerous watermarks (a group, in fact, includes two or more watermarks), the percentage of datable editions descends to just 25%.

Dating of editions through 'equivalence groups'. Editions involving at least three groups (tab. 6)

Months error	Actual percentage	Editions
-8 -4	13.64%	3
-4 0	31.82%	7
0 +4	40.91%	9
4 +8	13.64%	3
Total	-8 / +8 = 100%	22 = 25%

Tab. 6: Dating of editions through 'equivalence groups'. Editions involving at least three groups

Even so, broadly speaking it may be said that the results obtained are rather encouraging.

References

Atanasiu, Vlad (2004), 'Assessing Paper Origin and Quality by Means of Large-Scale Laid Lines Density Measurements', in Graziaplena, Rosella, with the assistance of Mark Livesey (eds), *Paper as a Medium of Cultural Heritage: Archaeology and Conservation*. 26th Congress IPH, Roma: Istituto centrale per la patologia del libro (Addenda, 5), 172–184.
Briquet, Charles-Moïse (1968 [1907]), *Les Filigranes. Dictionnaire Historique Des Marques Du Papier Dès Leur Apparition Vers 1282 Jusqu'en 1600. A Facsimile of the 1907 Edition with Supplementary Material Contributed by a Number of Scholars*, I–IV. Edited by Allan Stevenson, Amsterdam: Paper Publications Society, 1968, 1st edition, Genève, Paris: Alphonse Picard et fils.
Gerardy, Theo (1964), *Datieren mit Hilfe von Wasserzeichen*, Bückeburg: Grimme.
Grosse-Stoltenburg, Robert (1965), 'Beiträge zur Wasserzeichenforschung. Technische Varianten bei der Fertigung der Drahtform', in *Papiergeschichte*, 15: 73–79.
Ornato, Ezio / Busonero, Paola / Munafò, Paola F. / Storace, Maria Speranza (2000), 'Un procédé de datation des documents graphiques fondé sur la resemblance morphométrique entre les filigranes', in *Papiergeschichte als Hilfswissenschaft*, 23, Kongreß der Internationalen Arbeitsgemeinschaft der Papierhistoriker (Leipzig 30th August – 5th September 1996), Marburg: IPH Editions (IPH Congress Book, 11), 61–75.
Ornato, Ezio / Busonero, Paola / Munafò, Paola F. / Storace, Maria Speranza (2001), *La carta occidentale nel tardo medioevo*, I–II, Roma: Istituto centrale per la patologia del libro (Addenda 4).