Parameter analysis of the Glottal Area Waveform based on high-speed recordings within a synthetic larynx model

Aus der Phoniatrischen und Pädaudiologischen Abteilung

der Friedrich-Alexander-Universität Erlangen-Nürnberg

Der Medizinischen Fakultät

der Friedrich-Alexander-Universität

Erlangen-Nürnberg

Zur Erlangung des Doktorgrades Dr. med.

vorgelegt von

Yvonne Katrin Schulz

Als Dissertation genehmigt

von der Medizinischen Fakultät

der Friedrich-Alexander-Universität Erlangen-Nürnberg

Tag der mündlichen Prüfung: 14. Januar 2020

Vorsitzender des Promotionsorgans: Prof. Dr. med. Markus F. Neurath

Gutachter/in: PD Dr. Stefan Kniesburges

Prof. Dr. Michael Döllinger

Content

Abstract

Objectives Today's diagnostic methods in clinical practice considering dysphonia rely predominantly on the rich experience of the examiner. The use of objective parameters for the detailed diagnostics referring to the cause of the dysphonia in patients is not fully developed yet. This research faces the analysis and evaluation of specific synthetic vocal fold models regarding their correct representation of the physiological human phonation. Synthetic vocal fold models help to enable the research of the validation of objective parameter. Simulating the physiological human vocal folds, the synthetic vocal fold models offer some benefits in research in comparison to clinical studies with individual subjects: better accessibility (time management and examination methods), better reproducibility and no potential ethical conflicts. But although the synthetic vocal fold models should theoretically imitate the physiology as closely as possible unfortunately, they can only offer a simplification of the reality.

Design and methods This work is based on the experimental setup developed in "Fluid-Structure-Acoustic Interaction during Phonation in a Synthetic Larynx Model" by S. Kniesburges (2014). The most relevant parts of the experimental setup are the single- and the multi-layer models with three levels of stiffness of the silicone rubber. The dilatable diameter of the variable supraglottal channel and the insertable false vocal folds are important parts of the experimental setup as well. The influence of these components on the vocal fold models is part of the analysis. To compare the values of the objective parameters (fundamental frequency, shimmer, mean-jitter, jitter, harmonics-to-noise ratio [HNR], cepstral peak prominence [CPP] and maximum area declination [MADR]) of the synthetic vocal folds with human vocal folds, the data of various studies with subjects is examined. Further, the different single and multi-layer models in different configurations were analyzed to conclude the effect of the various components of the setup, as for example the false vocal folds or the supraglottal channel diameter, on the parameters and thus the vocal folds.

During the change of the setup the behavior of the objective parameters has been examined and compared with the values of the physiological data to establish a statement about the quality of the synthetic vocal folds.

Observation and results Overall, the synthetic vocal folds show great similarities to the physiology. The values of the parameters fundamental frequency, shimmer in percent, HNR and MADR are located within the same range as the physiological data. The values of the objective parameters of the single- and multi-layer models represent mainly the male physiological data, which is especially obvious within the data of the fundamental frequency.

Conclusions The total closure of the glottal gap during the phonation has a stabilizing effect on the oscillation of the vocal folds within both synthetic models, which can be seen in the reduction of the standard deviation of the fundamental frequency in comparison to no contact of the vocal folds during the phonation. In the analysis of the HNR data the influence during the change of the channel diameter seems larger in the multi-layer models than in the single-layer models. Additionally, the analysis of the MADR shows that the dilatation of the channel diameter leads to a faster closing of the glottal gap in nearly all models and oscillation modes. Another interesting point is shown in the analysis of the CPP data of the single- and multi-layer models: the insertion of the false ventricular folds raises the tonality and periodicity of the phonation. In conclusion, these synthetic vocal fold models represent the physiological human phonation in a valid way, which is seen in the similarities between the values of the objective parameters of the models and the subjects.

Zusammenfassung

Der Titel meiner Dissertation lautet: Parameteranalyse der Glottal Area Waveform basierend auf Hochgeschwindigkeitsaufnahmen von synthetischen Larynxmodellen.

Hintergrund und Ziele Die aktuellen diagnostischen Methoden im klinischen Alltag in Bezug auf Dysphonie beruhen fast ausschließlich auf dem Erfahrungsschatz des jeweiligen Untersuchers. Für eine detaillierte Diagnostik kann aktuell nicht auf verlässliche objektive Parameter zur Unterstützung des Untersuchers zurückgegriffen werden, da diese im Moment noch nicht ausgereift sind. In dieser Arbeit finden die Analyse und Beurteilung des künstlichen Kehlkopfmodells hinsichtlich einer korrekten Abbildung der Physiologie der menschlichen Phonation statt. Die Entwicklung von synthetischen Stimmlippenmodellen unterstützt den Versuch objektive Parameter valider zu machen, da diese im Gegensatz zu menschlichen Probanden einige Vorteile in der Forschung aufweisen: besserer Zugang (sowohl in Bezug auf das Zeitmanagement, als auch auf die Untersuchung), bessere Reproduzierbarkeit und keine ethischen Verpflichtungen gegenüber den Modellen. Doch um diese Vorteile nutzen zu können, müssen die synthetischen Modelle die menschlichen Stimmlippen so gut wie möglichen imitieren, wobei nur eine vereinfachte Version der Realität wiedergegeben werden kann.

Methoden Diese Arbeit bezieht sich auf den Aufbau und die Modelle aus „Fluid-Structure-Acoustic Interaction during Phonation in a Synthetic Larynx Model“ von S. Kniesburges (2014). Wichtig für den Versuchsaufbau sind die Einschicht- und Mehrschichtmodelle, welche in verschiedenen Steifheitsgraden untersucht werden. Ebenfalls entscheidend ist der weitenverstellbare supraglottale Kanal und die einsetzbaren Taschenfalten. Diese einzelnen Komponenten beeinflussen die Phonation der Stimmlippen maßgeblich. Um die objektiven Parameter (fundamental frequency, shimmer, mean-jitter, jitter, harmonics-to-noise ratio [HNR], cepstral peak prominence [CPP] und maximum area declination [MADR]) der synthetischen Modelle mit denen der Physiologie menschlicher Stimmlippen zu vergleichen, werden verschiedene physiologische Datensätze herangezogen. Zusätzlich werden die Modelle und deren unterschiedliche Konfigurationen untereinander verglichen, um den Effekt der einzelnen variablen Komponenten auf die synthetischen Stimmlippen zu erkennen.

Die unterschiedlichen Konstellationen der Modelle und deren daraus resultierenden Werte der objektiven Parameter sind im Anschluss mit denen der physiologischen Datensätze verglichen worden.

Ergebnisse Zusammengefasst bilden die beobachteten synthetischen Stimmlippenmodelle die Physiologie nicht in jeder Hinsicht nach, jedoch vor allem die Werte der fundamental frequency, shimmer in percent, HNR und MADR eine hohe Übereinstimmung zur den

physiologischen Datensätzen zeigen. Auffällig ist, dass die Werte der männlichen Datensätze stark mit den Werten der Ein- und Mehrschichtmodelle übereinstimmen. Vor allem in Bezug auf die fundamental frequency.

Schlussfolgerungen Der Kontakt der Stimmlippen beim kompletten Glottisschluss weist in beiden synthetischen Modellen eine stabilere Stimmlippenschwingung auf. Sichtbar wird dies durch die Verringerung der Standardabweichung der fundamental frequency im Vergleich zu keinem Kontakt der Stimmlippen bei der Phonation. Ebenfalls wird bei der Betrachtung der HNR deutlich, dass der Einfluss des supraglottalen Kanaldurchmessers auf die Mehrschichtmodelle größer ist als auf die Einschichtmodelle. Dazu kommt, dass die Vergrößerung des supraglottalen Kanaldurchmessers zu einem schnelleren Schluss der Stimmritze in fast allen Modellen und Schwingungsmodi führt. Des Weiteren weisen die Ein- und Mehrschichtmodellen beim Einsetzen der Taschenfalten eine erhöhte Tonalität und Periodizität auf. Die synthetischen Stimmlippenmodelle bilden die physiologische menschliche Phonation valide ab, was anhand der Übereinstimmungen der objektiven Parameter von Modellen und Probanden zu erkennen ist.

1 Introduction

1.1 Background and motivation

The human body is unique in its composition and function. Hence it takes much of time, investigation and ability to imitate the functional structure of the human larynx and to explore the process of phonation more precisely. The poor accessibility of the human vocal tract and larynx makes it difficult to examine the phonation process in detail. The subglottal area, for example, is only visible and accessible via the glottal gap. Because of this anatomic circumstance some measurement like the subglottal pressure are not possible in vivo.

To explore the process of phonation in depth, synthetic models provide a good alternative to in vivo measurements.

1.1.1 Anatomy

In the following, the basic anatomical structures and function of the pharyngeal and laryngeal region will be described shortly to provide a better overview of the process of phonation.

The original function of the larynx is to protect the lungs from aspiration by closing the airway [1]. This security mechanism is comprised of a three-stage safety valve that closes during swallowing: the epiglottis, the ventricular folds and the vocal folds.

The framework of the larynx is composed of cartilages, ligaments, tissue and muscles. The inside of this framework is covered with the mucous membrane.

The cricoid cartilage forms the transition to the trachea and is attached to the thyroid cartilage, which surrounds the larynx from ventral and leaves a gap that consists of soft dorsal tissue. The two arytenoid cartilages build the suspension for the vocal folds and enable the muscularly controlled movement of the vocal folds.

The hyoid bone serves as an insertion to the pharyngeal muscles and is attached to the thyroid cartilage by the thyrohyoid membrane. Inside the larynx, the epiglottal cartilage forms the foundation of the epiglottis. The epiglottis seals the airway, so that neither solid, nor liquid material may enter it.

In addition to enabling the production of sound, the oral cavity, nasal cavity and pharynx cavity are in parts of crucial importance for the process of food consumption, too. All these terms – including the larynx – are subsumed under the term vocal tract. Its dimensions are illustrated in Figure 1.

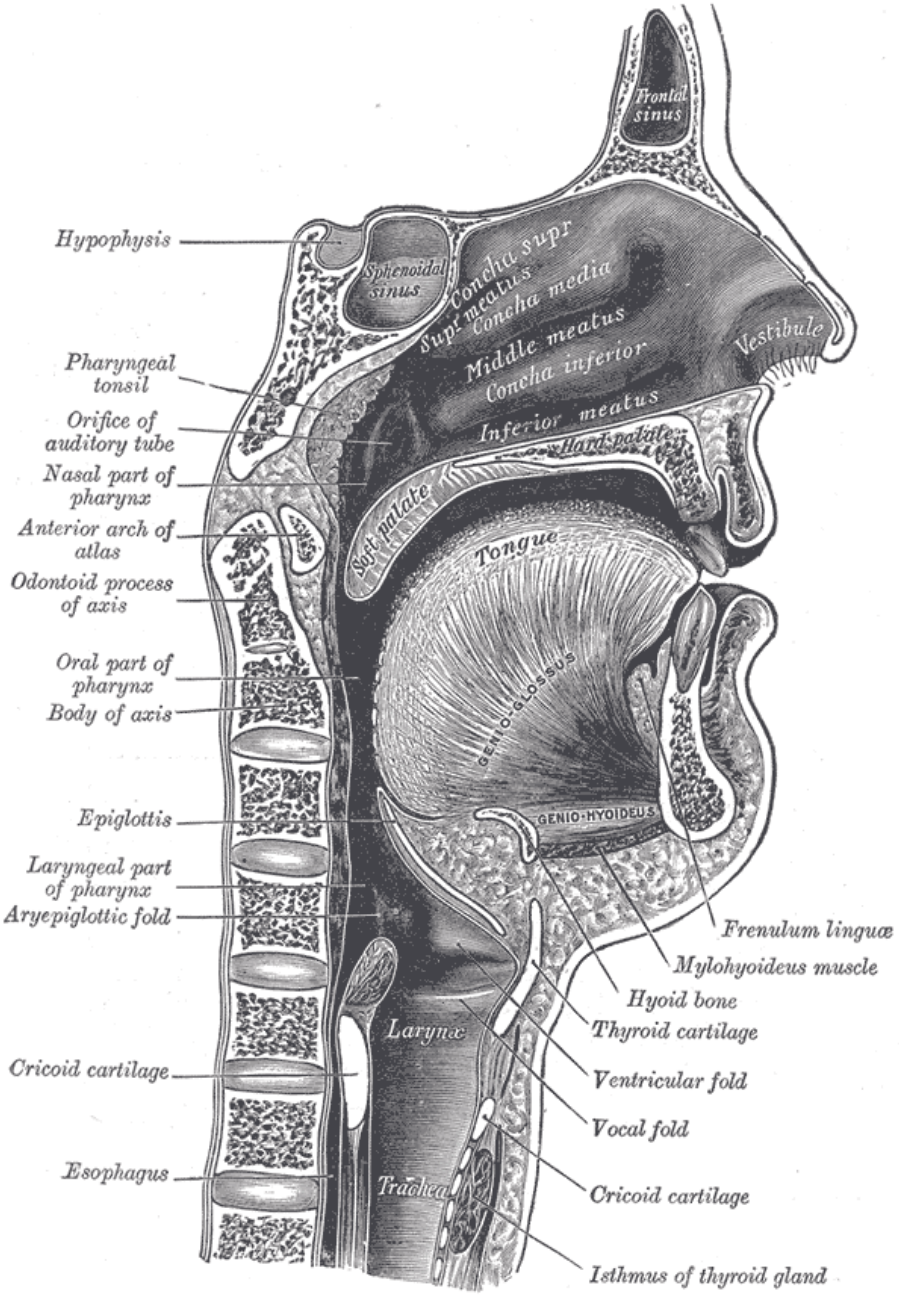


Figure 1 - Sagittal section of nose mouth, pharynx, and larynx, taken from [2].

The larynx is topographically divided into the subglottal, glottal and supraglottal area, see Figure 2.

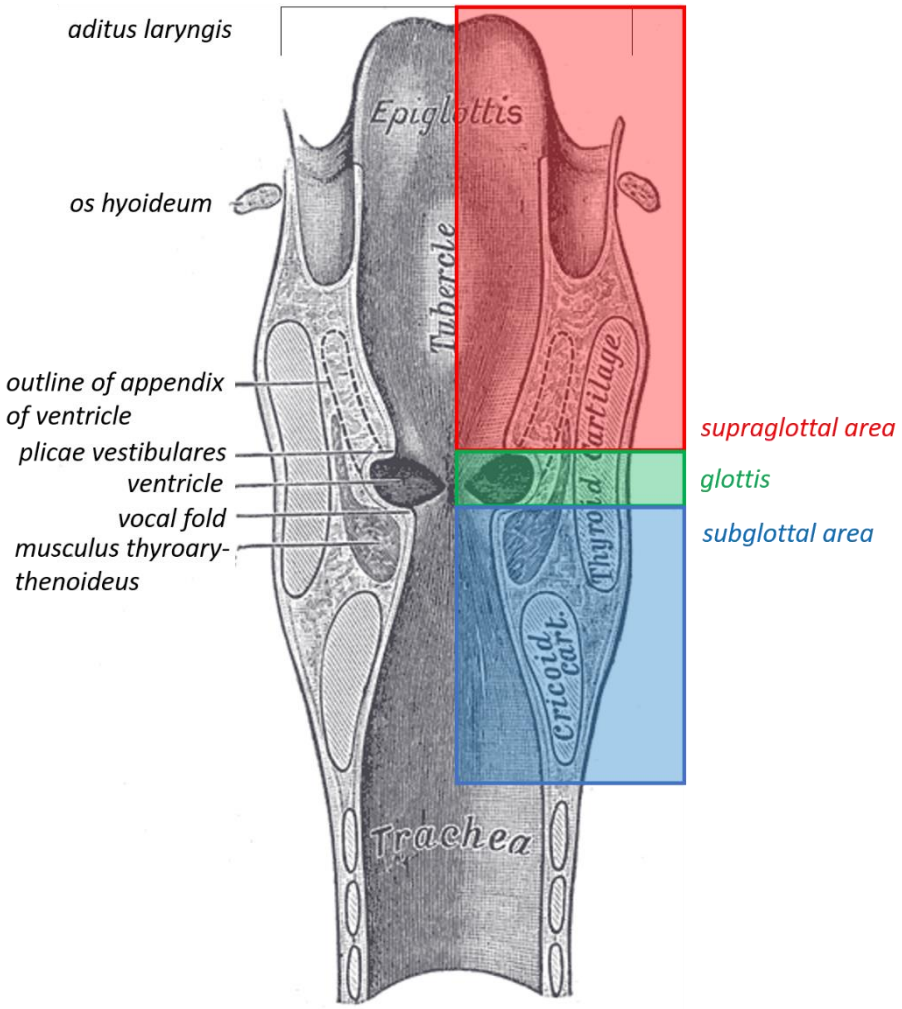


Figure 2 - Coronal section of larynx and upper part of trachea based on [2].

The supraglottal area is defined by the *adtius larnygis* and the *plicae vestibulares*, also known as the false vocal folds. Inside this area the epiglottis is located. The gap between the vocal folds is the glottis, which is the doorway through the trachea into the lung, see Figure 3. The subglottal area extends from the vocal folds to the lower rim of the cricoid. [3, p. 204]

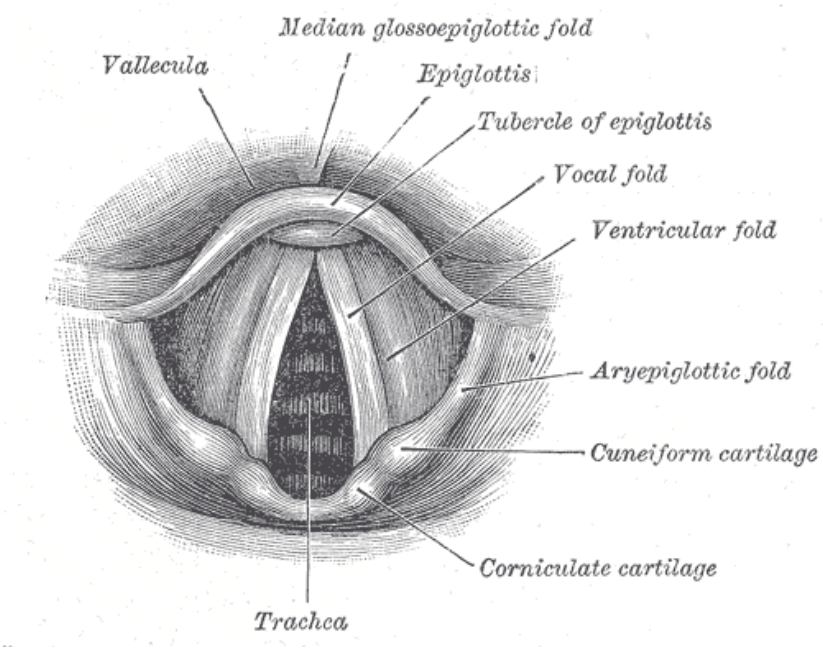


Figure 3 - Laryngoscopic view of interior of larynx taken from [2].

The vocal folds consist of four parts: epithelium, ligament-cover (with a superficial layer, intermediate layer and a deep layer) and body. The vocal folds are covered by squamous epithelium to protect them from physical damage. Loose conjunctive tissue, called superficial layer, is attached to their bottom side and is followed by the ligament. In the deeper layers of the vocal fold tissue the *musculus vocalis* adheres to the ligament and makes up the body of the vocal folds (Figure 4). [4]

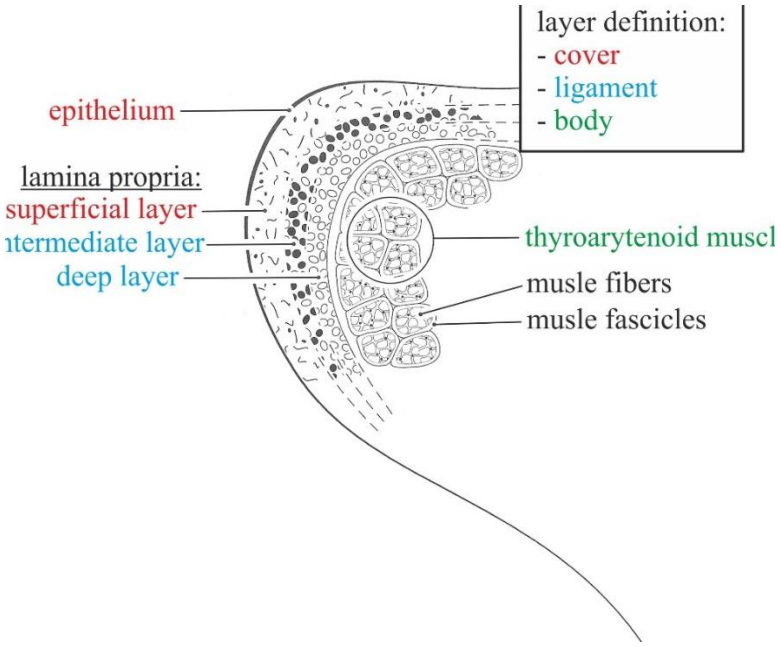


Figure 4 – Schematic cross-section of the vocal fold taken from [5, p. 6].

Figure 5 shows the muscles which enable human voice production. These muscles perform the pre-phonatory positioning motion of the vocal folds and affect their tension and the dimension of the glottal area.

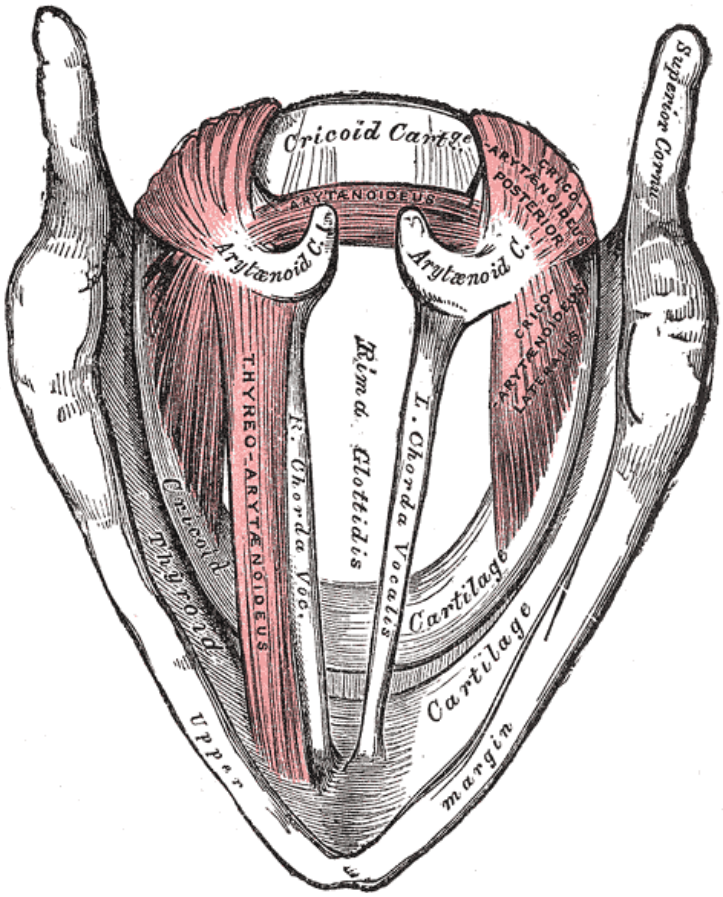


Figure 5 - View on the glottis and the vocal folds – muscles of the larynx exposed taken from [2].

A group of muscles is responsible for the glottal closure by sealing the gap between the two vocal folds which adjusts the glottal gap to its respective necessary function, like breathing, whispering or speaking. Another group of muscles is responsible for adjustment of the vocal folds' tension.

The muscles which will be discussed subsequently are all located inside the larynx. The opening of the gap is performed by the *musculus cricoarytaenoideus posterior* and partwise by the *musculus cricoarytaenoideus lateralis*. Both originate from the cricoid and rise to the arytenoid cartilage.

The three muscles - *musculus arytaenoideus obliquus*, *musculus arytaenoideus transversus* and *musculus thyroarytaenoideus* - close the glottis by pulling the arytenoid cartilages together in transverse plane and pushing them ventrally. This movement is comprises the rotation and the

translation, which is proceeded by the *articulatio cricoarytaenoidea* and is called rocking motion.

The vocal folds enter at the arytenoid cartilage. The ventral insertion of the vocal folds is fixed at the thyroid cartilage. This fixation leads to a motion in the transverse plane. The movements of the arytenoid cartilages towards each other closes the glottal gap. The movement towards the opposite direction opens the glottal gap. Because of the stiff ventral insertion, the open glottal gap has a triangular shape.

Furthermore, the *musculus vocalis*, which runs inside the vocal folds, is responsible for the tensing, thickening and contracting of the vocal folds [6].

1.1.2 Physiology

The oral and pharyngeal cavities main function is the intake of food and drinks as well as facilitating air exchange in the lungs. Fresh air is filtered, humidified and preheated in the nasal and oral cavity. Additionally, the modulation of the sound is also performed in the vocal tract, which creates the characteristic signal of sound.

The production of the human voice is a very complex process with many different components and participants. The vocal folds produce the basic sound of voice production by oscillating in a specific frequency. This oscillation is caused by the airflow from the lungs through the trachea and glottis (in between the vocal folds).

The position of the vocal fold is crucial for voice production: the laryngeal muscles pull the vocal folds together and thereby close the glottal gap almost completely. The subglottal pressure opens the vocal folds and the flow between them leads to the oscillation of the vocal folds. Consequently, the flow speed in this narrow glottal gap is higher than the flow speed in the subjacent trachea. The laryngeal muscles pull the vocal folds together and through Bernoulli's principle[1] the air pressure decreases which leads to a reduction of the glottal gap which also pulls the vocal folds together until the total closure of the glottis is reached. When the glottis is closed, the air flow is interrupted. The vocal folds are pressed apart again by the increased subglottal pressure, which is built up by the lungs for voice or speech production. [7, p. 754]

The typical movement of the vocal folds during the opening and the closure of the glottal gap sets the surrounding air in motion. This oscillation of the vocal folds takes place in three dimensions: in the transversal plane the wave starts at the lateral ends during the opening of the glottal gap and is transferred to the ventral parts of the vocal folds. Meanwhile, in the sagittal

[1] Bernoulli´s principle: total pressure equals static pressure plus dynamic pressure. [60, p. 100]

plane the vocal folds move up and down, which generates a wave-like motion in the vocal folds. Additionally, the lateral edges of the mucosa of the vocal folds slide over the edge of the vocal folds and cause the mucosal wave motion. [8, pp. 19-20]

The variability of voice allows us to communicate or sing. The tone pitch is adjustable trough the change of length and tenderness of the vocal folds. Due to this adjustment the oscillating length of the vocal folds changes and the partial tone alters in frequency and amplitude. For example, the tension of the vocal folds rises to produce a high tone with a higher frequency and a smaller amplitude. [9, p. 85]

The loudness of the voice is adaptable by the subglottal pressure and the adjustment of the respiratory flow speed. The higher the subglottal pressure is and therefore the faster the respiratory flow gets, the louder is the produced the voice. The maximum duration of phonation depends on various factors such as flow speed and vital capacity or age and gender. [9, p. 85]

The produced basic voice signal is carried through the vocal tract, which is composed of the supraglottal area, the nasopharynx and the oral cavity. The vocal tract forms a resonance body in which the sound and its respective frequency is modulated by the increase or decrease of the partial tones [8, pp. 18-19]. The voice signal is filtered through the absorption of some partial tones due to the mismatch between the frequency of the produced sound and the shape of the vocal tract. The shape of the vocal tract is adaptable by the muscle tube, which builds the pharynx. During the contraction of the pharyngeal musculature the tube gets narrow and the production of tones can be affected. It is also different from individual to individual by the different sizes of the specific anatomical body parts. [9, p. 21]

1.2 State of the art

To get an overview of the various larynx models and their possibilities and limits, these model types will be discussed briefly in the following chapter.

1.2.1 Various model types

There are three main reasons for further development of synthetic larynx models. First, the lack of available excised human larynx and their short usability [10]; second, the usage of animal larynx serves the purpose of fundamental research, but the models are only similar in function and composition, not identical. The histological structure of the animal’s larynx differs from the human larynx in its number of layers as well as its collagen types. Also, the animal anatomic structure is different from its human match, displayed in the size of the false vocal folds for example [11, p. 9]. Thirdly the measurements on living human beings are limited by ethical reasons or problems of access [12, p. 305].

Experimental and numeric larynx models are often of great interest to researchers, who are trying to measure the fluid structure interaction of the vocal folds. Since the correlation between flow and oscillation of the vocal folds is only measurable with the help of these models. The structural oscillation of the human vocal folds is easily detected by endoscopy, but it is not possible to measure the flow inside the larynx in a living human. This last aspect explains the importance of larynx models of any kind.

To bypass this problem different methods are used [5, pp. 16-17]:

1. Excised larynx models:

These models are taken out of human bodies or animals. The measurements are quite difficult, and the missing muscle activity impedes the realistic display of the human physiology. Additionally, they remain viable only for a short time due to tissue degeneration, resulting in a poor reproducibility of measurements. [10, p. 286–304] Ex vivo rabbit larynges are used to examine their behavior and measurements of objective parameter for example in Maytag et al. (2013) [13]. Other studies explored the effect of supraglottal anatomic structures with the help of ex vivo canine larynges like Alipour and Finnegan (2013) [14].

2. Synthetic larynx models:

The realistic imitation of the human larynx made of silicone rubber [15] or other components [16] enables a good reproducibility of measurements. Also, various kinds of vocal fold models can be used and compared. The time limit of the synthetic larynx measurements is longer than for the excised larynx models and is limited only by the mechanical load and their effect on the used material.

However, synthetic models are limited by their restricted function. The adduction or abduction, as well as the pre-stress of the vocal folds is still content of scientific development. Though, the vibration characteristics of the synthetic larynx models are similar to the movements of the human vocal folds. [12]

3. Numeric computer models:

To simulate the process of phonation in a computer model numerically, extensive computation effort is necessary [17]. For example, it takes days to calculate a small amount of cycles of vibration. For these complex computations, massively parallel high-performance computers are required. Hence, two-dimensional numeric models serve as a workaround [18]. This method still needs further development [17].

1.2.2 Different model types of synthetic models

The development process of synthetic larynx models started in 1957 with van den Berg [19]. According to Kniesburges et al. (2011), todays in vitro experimental investigation in synthetic laryngeal models can be divided into three groups [12]:

1. Static models

The first ever documented static model was created by Wedel (1930) and it is made from a plaster cast of a canine larynx [20]. As the name suggests, the static models include a ridge vocal fold model within a stiff flow channel and a constant or variable flow rate [20]. During the development of the static models, Scherer et al. established the M5 model geometry, which many other researchers have adapted [21].

The static models are useful for aerodynamic studies, because of their simplified representation of the reality which leads to a focus only on the aerodynamics in extremely controlled and reproducible conditions [12].

2. Externally driven models

At the end of the 20th century, most efforts were put into externally driven models. The vocal fold movement is in focus of these kind of studies. There are two types of externally driven models: uniform-phase driven models and non-uniform-phase driven models. These two are distinguished by their vocal fold model motion. The uniform-phase driven model allows a motion in medial-lateral direction [22], whereas the non-uniform-phase driven model can move in medial-lateral and superior-inferior direction, which leads to the typical human convergent-divergent motion as shown in the experiment by Triep et al. [23].

The externally driven models are mainly used to incorporate the vocal fold motion in aerodynamic measurements and to provide data for computer analysis [12].

3. Self-oscillating models

In the last ten to fifteen years, the flow-induced motion of the vocal folds has been included. The models are made of polyurethane or silicone rubber and can be compared to in vivo

experiments. The structural behavior, aerodynamics and the acoustic field are the respective objects of interest. [12]

The third model type, the self-oscillating model, is discussed in more detail in the context of this dissertation and can be divided into two varying model types [2]:

a. Single-layer models

The single-layer vocal fold models are completely made from a homogeneous silicone rubber that are formed mainly based on the M5 model of Scherer et al. [21].

Their flow-induced vibrations are an important requirement for further research regarding the function of voice production. The air is flowing through the subglottal channel and the vocal folds.

Thomson et al. designed a hemilarynx in life-size and later a complete laryngeal setup which show the change of the glottal profile from convergent to divergent shape during the oscillation [24].

High-speed cameras were used to record the oscillation behavior [25]. Those models were also used in the setup for the experiments performed described in [5] which are described in chapter 2.

With this type of models, especially "[…] the structural dynamics of the vocal folds, the fluid flow in the subglottal region, and the acoustic signal spread to the far field" [12, p. 318] are the subject of research within the single-layer models.

b. Multi-layer models

Multi-layer models are beeing developed in research for a more realistic composition of the vocal folds, to investigate the effect of the thickness of the different layers and their flow-behavior [26]. Drechsel and Thomson (2008) [27] used the template of the shape of the M5 models of Scherer et al. (2001) [21] to build up a model with two layers. These synthetic vocal folds are placed in a synthetic larynx with three different supraglottal cases: open jet, simplified vocal tract formation and vocal tract formation with FVF. Additionally, the position of the supraglottal tube could be changed to observe the behavior of the supraglottal flow field [27]. [12]

[2] This is also part of this dissertation trough the comparison of the single and multi-layer models in chapter 3.1

They also build multi-layer models composed of more than two layers, such as the four-layer vocal fold model, which is composed of an epithelia layer, lamina propria layer, ligament and body layer [28]. The multi-layer models used in experiment in [5] have three layers. [12]

To analyze the features of multi-layer models in comparison to the human physiology, new models based on MRI data are built [29] and the M5 model of Scherer et al. (2001) [21] is replaced in some cases by a more realistic model setup [30]. The mucosal wave motion is in focus of research. [12]

c. Membranous models

Membranous vocal fold models are made from a thin silicone rubber membrane forming the contour of the vocal folds. Inside, these vocal fold models are filled with fluids or air [31]. The subject of the research is also the mucosal wave motion and phonation threshold pressure [5]. [12]

1.2.3 Digital high-speed visualization

The usage of digital high-speed cameras in research and clinical praxis offers the possibility of an objective analysis of the vocal folds and their potential dysfunction. It is thereby "(…) the beginning of evidence based diagnostics in endoscopic voice diagnostics." [19, p. 101]. Other measurement methods like videostroboscopy[3] are limited to diagnostics in voice disorders [32]. The human vocal folds oscillate at a frequency of 100 to 300 Hz, which cannot be captured in total with normal digital cameras that normally provide framerates below 50 Hz [19]. This leads to a loss of information in cases of voice disorders with highly aperiodic vocal fold oscillations owing to the low frame rate of normal cameras in videostroboscopes.

Recordings captured by the videostroboscopes only show the complete cycle of the vocal fold movements in completly periodically cycles. In not perfectly periodical cycles only the phonation onset and a deviation from the periodicity of the vocal fold oscillation can be detected. This is because of the shortening of the oscillation cycles through a quasi-slow motion. The oscillation cycles must be perfectly periodic to allow an analysis of the oscillation of the vocal folds with the videostroboscope. If the oscillation is aperiodical, the display by the videostroboscope shows only a chaotic sequence of pictures. With a high-speed videoendoscope one complete cycle of the vocal fold movement is recorded. More pathologies such as left-right asymmetry or irregular oscillations can be detected. [19]

[3] Flashing light in constant intervals that are slightly longer than a multiple of the period length of the vocal fold oscillation decelerate the movement of the vocal folds optical.

Still the videostroboscopy is the gold standard [33] and the high-speed videoendoscopy (in the following refered to as HSV) is only an additional diagnostic method in the daily clinical routine [34].

The challenge of the high-speed videoendoscopy is the great effort of analyzing the HSV data. Therefore, an image processing algorithm for the objective and quantitative diagnostic is nessesary [35]. Also, the purchasing costs of HSV are very high.

The first usage of HSV started in the late 1930s and has been used for irregular or aperiodical voice disorders [36]. The problem, at this time, has been the evaluation of the analog images. This problem is solved by introducing the digital image processing techniques [37].

Today, the interpretation of the videos recorded by the HSV is analysed by physicians. They examine the videos and interpret them based on their experience. Timcke, von Leden and Moore started in 1958 with the objectification of the HSV [38].

For current work, the videos were segmented by “Glottis Analysis tools” (GAT). The glottal area in the video is detected semi-automatically as function of time. The program detects the contrast change between the black glottis and the brighter tissue and therewith the glottal area. The base of the analysis is the glottal area waveform (GAW), which represents the glottal area as function of time [39]. The next step is the mathematical analysis of the GAW in order to obtain objective information on the vibration of the vocal folds [19].

The HSV is mainly used in research of the voice function [40]. But to attain the standard of evidence-based medicine in voice diagnostics the next years of research could cause a change in the daily clinical routine [41].

This might be because of the possibility of objective interpretation of the HSV videos by objective parameters calculated by standardized programs like GAT.

Furthermore, the HSV is recording a direct reflection of the motion inside the larynx and does not need to rely on acoustic signals as the videostroboscopy. For this reason, and the better quality of the recordings done by the HSV, it could significantly ease and improve the diagnosis in a clinical setting [42].

The better quality of the HSV videos directs to a more accurate prediction of the characteristic vibrations of the vocal fold tissue, especially regarding the mucosal wave and amplitude of vibration diagnostics [43]. n

All in all, the usage of HSV is not yet practicable in daily clinical routine due to the quantity of data that needs to be processed [44, p. 261].

2 Experimental setup and analysis methods

This dissertation is based on the experimental work by Stefan Kniesburges, described in his PhD-thesis “Fluid-Structure-Acoustic Interaction during Phonation in a Synthetic Larynx Model” in 2014 [5]. The following description of the experimental setup as well as the performed measurements constitutes a summary of his PhD-Thesis [5].

The experimental setup was composed of a synthetic larynx with two different vocal fold models placed in a life-size mounting device and an adjustable supraglottal channel. The vocal fold oscillations were recorded by a high-speed camera. The high-speed camera detected the vocal fold movements with a maximum resolution of 1024 x 1024 pixels and a frame rate of up to 5400 frames per seconds (fps) [5, p. 59]. With these video files and the program GAT, it was possible to segment the glottis and to analyze several parameters based on the GAW.

2.1 Synthetic larynx model

2.1.1 Complete channel

The entire setup can be divided into 2 parts: the measurement setup and the synthetic larynx models

The synthetic larynx was built in a human length-scale representation of a human larynx. The whole setup can be seen in the following.

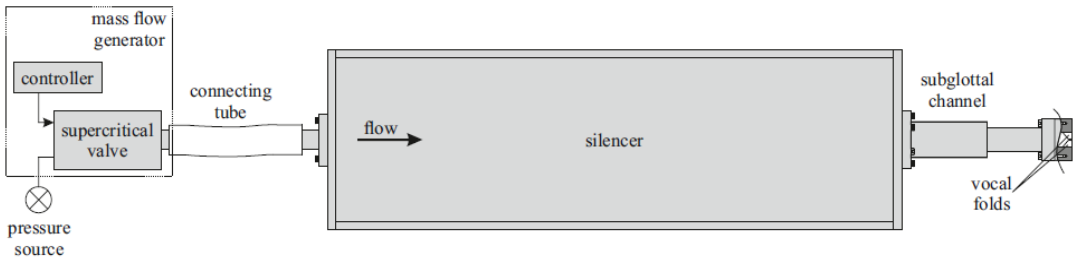


Figure 6 - Schematic of the test channel up to the vocal folds taken from [5].

Figure 6 displays the setup of the whole test channel schematically including mass flow generator (consisting of the controller and the supercritical valve), the pressure source, the connecting tube which connected the mass flow generator to the silencer. The silencer was connected to the subglottal channel leading to the vocal folds.

The electronically controlled mass flow generator produced the air flow with adjustable flow rate from 0 to 180l/min [5, p. 23]. The silencer was subdivided by a flow passed absorbing body that was located in the center of the silencer. The aim of the silencer was to minimize acoustic disturbance, which was generated by the valve of the mass flow generator [45].

It was possible to force a specific amount of flow rate through the subglottal channel and therefore through the glottis into the supraglottal channel. The high-speed camera visualized the movements of the vocal folds induced by the air flow through in the supraglottal channel. The measurement setup is illustrated in Figure 7.

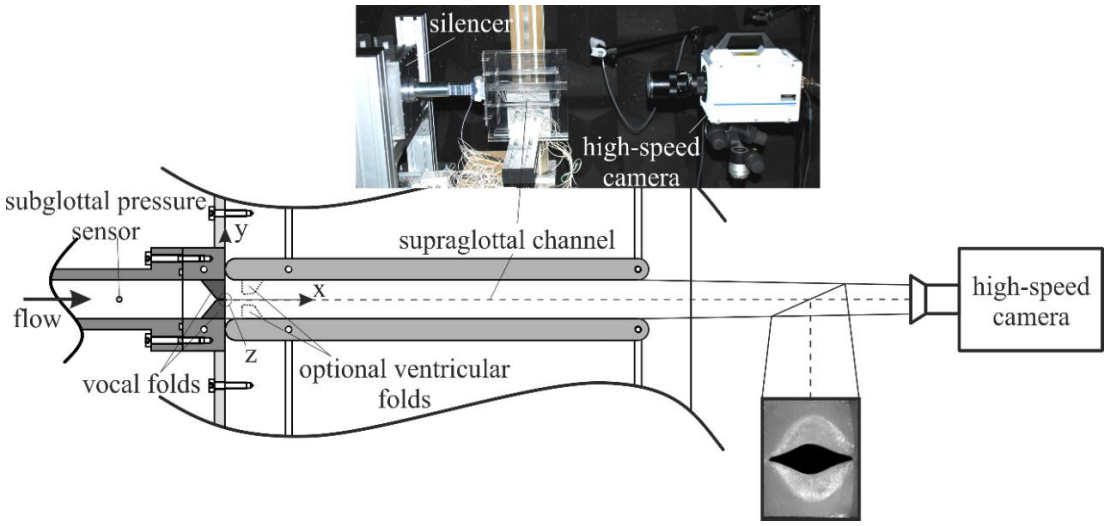


Figure 7 - Experimental setup for the visualization of the structural vibration of the synthetic vocal fold models. The figure is taken from [5].

A Kulite XCQ-093-SG sensor measured the subglottal pressure during the recording of the high-speed camera. 500 images in 2 seconds were captured during the recording of the vocal fold movement and the data of the subglottal pressure measurement was captured synchronously.

The global dimension of the subglottal channel was based on human anatomy. The size of the rectangular cross-section of the subglottal channel was 0.015m x 0.018m. The length of the channel was 0.19m. With this setup it was possible to induce aerodynamically driven vibrations of vocal fold models. [5, p. 26]

The synthetic vocal folds were installed within the mounting device and attached to the supraglottal channel. This is illustrated in Figure 8.

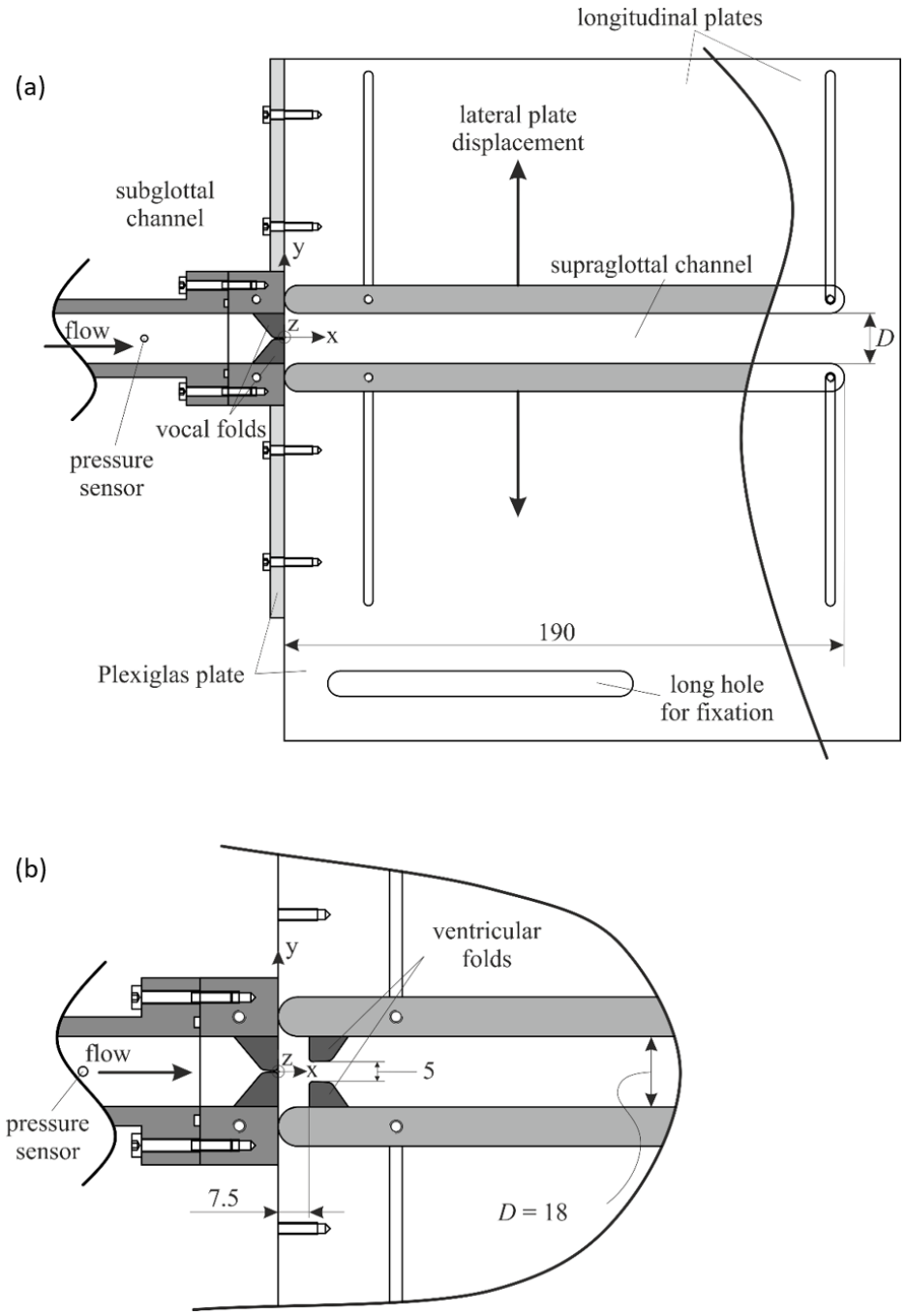


Figure 8 - Schematics of the configurations within the supraglottal flow region including the supraglottal channel. Filled faces represent cut views of the channel components. As the front longitudinal plate is made of glass it is sketched transparent to uncut areas: (a) Complete supraglottal channel without FVF (b) Zoomed image of the immediate supraglottal region including the FVF. All dimension in [mm]. The figure is taken from [5].

The supraglottal channel consisted of four Plexiglas plates, two large plates for limiting the channel in longitudinal and two small plates for limiting in lateral direction which can be moved in medio-lateral directions. This was designed to observe the impact of the supraglottal channel dimensions on the vocal fold vibrations. Also, it led to the option of changing the supraglottal channel volume by altering the cross-section of the supraglottal flow region. This will be explained in more detail in chapter 2.1.1.

The two longitudinal plates kept a constant longitudinal diameter of 15mm, while the lateral plates can be shifted with diameters between D 18mm - 78mm. Additionally, the supraglottal channel can be removed completely. The total length of the supraglottal channel was 190mm. [5, p. 43]

2.1.2 Vocal fold models

Two different types of vocal fold models in human-life size were placed in the measurement setup. The geometry of all vocal fold models analyzed in this thesis were based on the M5 model [21].

1. The single-layer models consisted of one layer of homogeneous two-component silicone rubber [5, p. 37]. They are named 114 and 113 in the following. The Young´s modulus of elasticity [4] of the 114 vocal fold model was 2.5 kPa, while the Young´s modulus of elasticity of the 113 vocal fold model was 4.4 kPa.
2. The multi-layer model had a more realistic structure compared with human anatomy. It was composed of a body, a ligament and a cover layer each with a different Youngs modulus [5]. The different stiffness values of the parts were achieved by adding a specific amount of silicone-thinner to the liquid silicone rubber mixture [5, p. 37]. They are named C115 and C114 in the following. The multi-layer models C114 and C115 varied only in the Young´s modulus of elasticity of the cover. Here the C114 model had a more elastic cover than the C115 model.

[4] Young´s modulus of elasticity: "The ratio of normal stress to corresponding strain below the proportional limit of a material" [59, p. 489]The higher E, the less elastic is the material.

An overview of the analyzed models with their Young´s moduli of elasticity are shown in the Table 1.

Vocal fold model	Mixing ratio part A – part B- thinner	Static Young´s modulus of elasticity *E* / kPa
113	1-1-3	4.4
114	1-1-4	2.5
C114	body: 1-1-2 ligament: 1-1-0 cover: 1-1-4	8.2 60 2.5
C115	body: 1-1-2 ligament: 1-1-0 cover: 1-1-5	8.2 60 1.52

Table 1 - Notation, volumetric mixing ratios and static Young´s moduli of the silicone mixture used for the vocal fold models according to [5, p. 39].

The structure of the single-layer and multi-layer models is illustrated in Figure 9. The geometry of the vocal fold models mimicked the human vocal folds, while the inner structure differed in between the two different model types. The structure of the multi-layer model was better adapted to the inner structure of the human vocal folds, as shown in chapter 1.1.1.

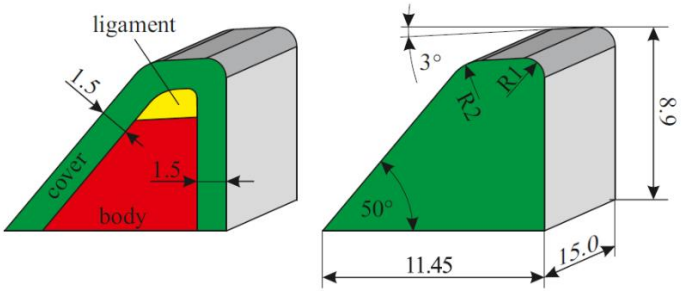


Figure 9 - Schematic of the single- (left) and multi-layer model (right) of a human vocal fold. All dimensions in [mm]. The figure is taken from [5, p. 37].

During the experiments, two different modes of oscillation were detected. In oscillation mode one (M1), the vocal folds oscillated periodically but never closed the gap between the vocal folds. Conversely, in oscillation mode two (M2), the vocal folds came in contact during the periodical glottis closure. Only the longitudinal ends of the glottis had no contact during glottis closure. M2 lead to a periodical interruption of the glottal flow. The two modes are illustrated in Figure 10 and Figure 11. [5]

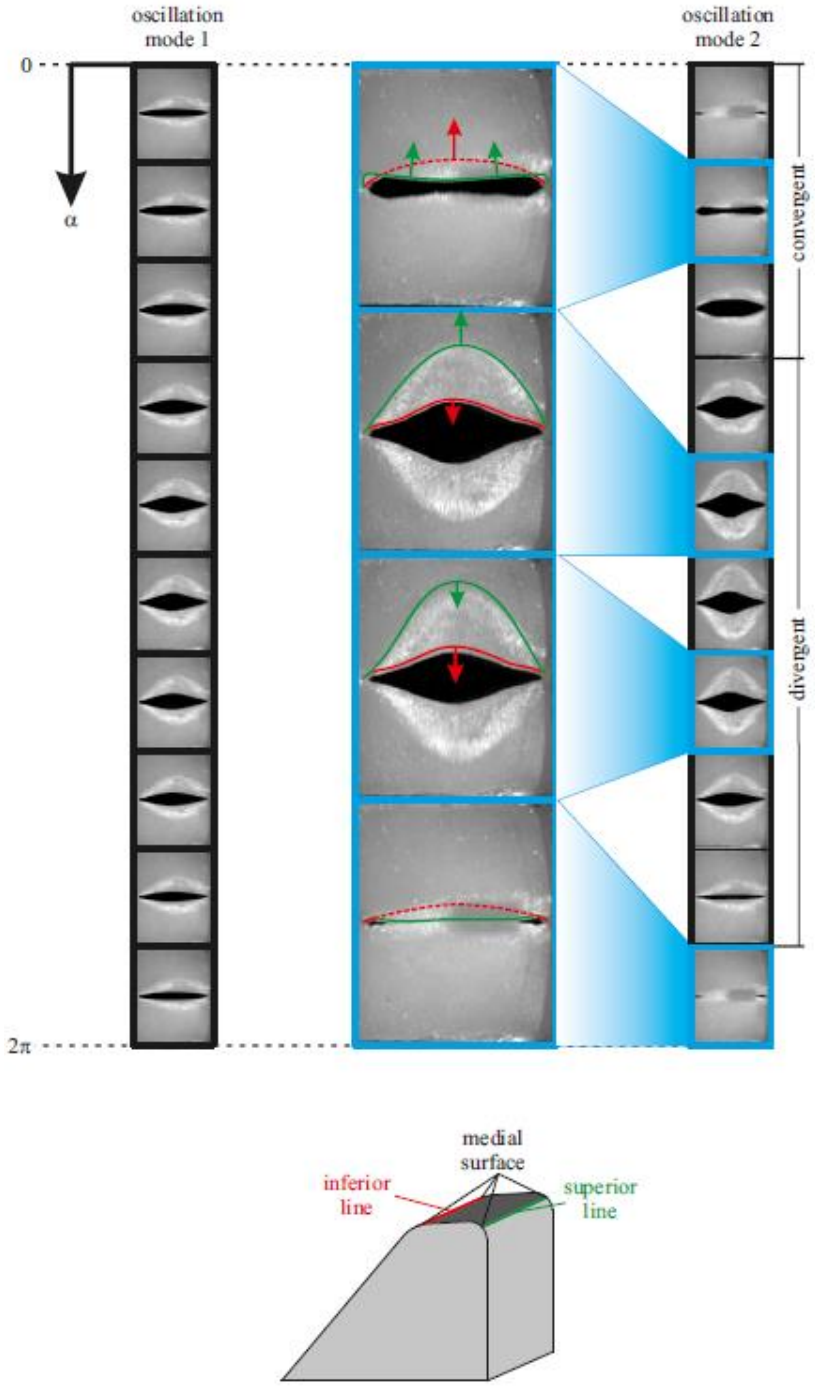


Figure 10 - Top view on the vocal folds of the single layer model 113 during one oscillation cycle. On the left side M1 and on the right-side M2. Underneath the schematic setup of a synthetic vocal fold is illustrated. The figure is taken from [5, pp. 62, 63].

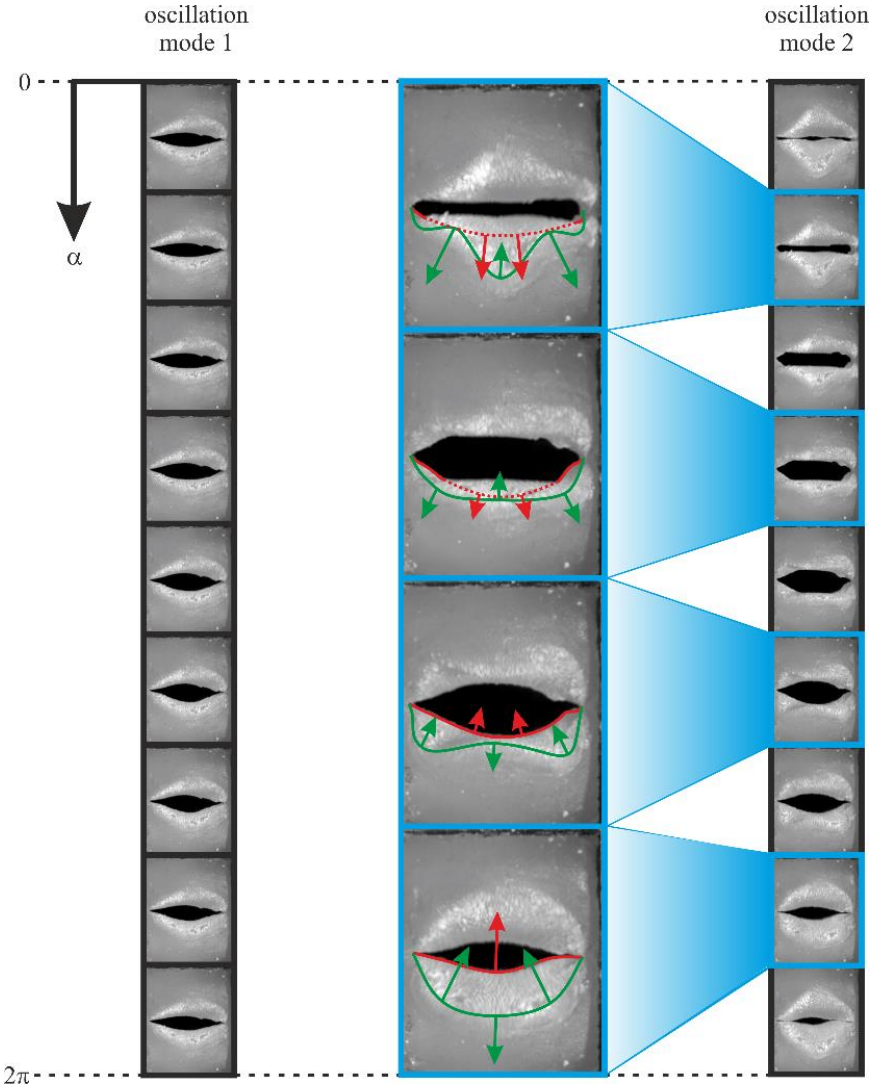


Figure 11 - Top view on the vocal folds of the multi-layer model C114 during one oscillation cycle. On the left side M1 and on the right-side M2. The figure is taken from [5, p. 65].

2.1.3 Ventricular fold models

The ventricular folds, made of hard plastic, can be installed optionally cranial to the vocal folds inside the supraglottal channel. Between the ventricular folds and the vocal folds, there was a constant distance of 7.5mm. [5, p. 43]

In the following text, false vocal folds will be denoted as FVF, which is a synonym for the false ventricular folds also.

The FVFs did not vibrate during the process of phonation [5, p. 43]. The size of the FVFs themselves was 10mm x 15mm at the base, but its shape was similar to the geometry of the vocal fold model, which can be seen in the following Figure 12.

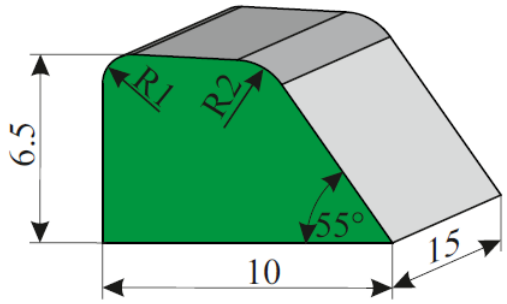


Figure 12 - Geometric dimensions of the FVF. All dimensions in [mm]. The figure is taken from [5, p. 43].

2.2 Description of supraglottal channel configurations

Eight different configurations were assigned based on the presence of FVF and the diameter of the supraglottal channel.

To get a better overview the different configuration numbers are listed in Table 2.

Configuration number	Properties		
	supraglottal channel	*D* / mm	ventricular folds
1	x	18	
2	x	28	
3	x	38	
4	x	48	
5	x	58	
6	x	78	
7		∞	
8	x	18	x

Table 2 - List of supraglottal channel configurations. Included features (supraglottal channel and ventricular fold models) are marked with x. The table is based on the table from [5, p. 44].

Every model configuration was measured in M1 and M2.

2.3 Measurement strategy

To visualize the movement of the vocal folds a high-speed camera with 512x512 pixels was installed. The high-speed camera was placed 300mm downstream of the supraglottal channel exit. To avoid an interference with the supraglottal flow this position was chosen. The high-speed camera recorded the oscillation of the vocal fold with a frame rate of 4000fps. [5, p. 59]

The variable flow-rate changed the subglottal pressure in every configuration and vocal fold model constellation. The subglottal pressure was measured constantly to provide the required frequency of the measurement setup [5, p. 28]. The resonant frequency varied during the change

of the supraglottal volume and the several model types. Every channel configuration and model type required an adjustment of the air pressure. [5]

Configuration number	113				114			
	M1		M2		M1		M2	
	flow rate [l/min]	Psub [Pa]	flow rate [l/min]	Psub [Pa]	flow rate [l/min]	Psub [Pa]	flow rate [l/min]	Psub [Pa]
1	100	3152	100	3251	105	2084	105	2302
2	110	3462	110	3517	105	2270	110	2408
3	110	3555	115	3559	110	2332	115	2527
4	110	3656	110	3736	110	2449	120	2686
5	120	3869	130	4094	115	2542	125	2789
6	130	4269	140	4528	115	2790	125	3028
7	130	4215	145	4400	115	2549	135	2901
8	75	2103	75	2449	70	1301	75	1579

Configuration number	C114				C115			
	M1		M2		M1		M2	
	flow rate [l/min]	Psub [Pa]	flow rate [l/min]	Psub [Pa]	flow rate [l/min]	Psub [Pa]	flow rate [l/min]	Psub [Pa]
1	90	3394	90	3809	66	2050	80	2464
2	100	3925	100	4332	75	2509	80	2710
3	105	4059	115	4771	75	2474	90	2932
4	105	4224	115	4850	75	2619	85	2907
5	90	3479	105	3775	75	2192	90	2526
6	95	3596	120	4435	80	2292	95	2687
7	120	4065			85	2446		
8			85	3249	65	1427	70	2193

Table 3 - Flow rate and mean subglottal pressure during M1 and M2 in all channel configurations. Top table: Parameters for vocal fold models 113 and 114. Bottom table: Parameters for C114 and C115. Grey fields replace the missing high-speed videos and their data. The values are taken from [5].

The flow-rate had a variation of 65-180 l/min provided by the mass flow generator [5, pp. 30-31] and it was adjusted to the vocal folds and the different supraglottal channel configurations. The key data in Table 3 shows the values of the mean flow-rate in M1 and M2. This key data displays the required effort for phonation of the models. [5]

2.4 Analysis of high-speed videos

Based on the experimental data described above, in this work an evaluation of the of the vocal folds oscillations is performed with the software GAT. This software was developed to analyze endoscopic videos of the human vocal folds oscillations taken from patients and subjects objectively.

2.4.1 "Glottis Analysis Tool" analysis

The recorded high-speed videos are imported into the GAT program, which was developed by the Institute of Phoniatrics and Paediatric Audiology at the University Hospital Erlangen.[5]

The video files are segmented in GAT by determining the frames of the glottal gap and its descriptive area, as shown in light blue in the following Figure 13. Seed points (green crosses in Figure 13) are inserted in the glottis to define the starting point for the region growing methods. This method determines the glottal area in all 500 video frames resulting in the GAW which is function of time. [46]

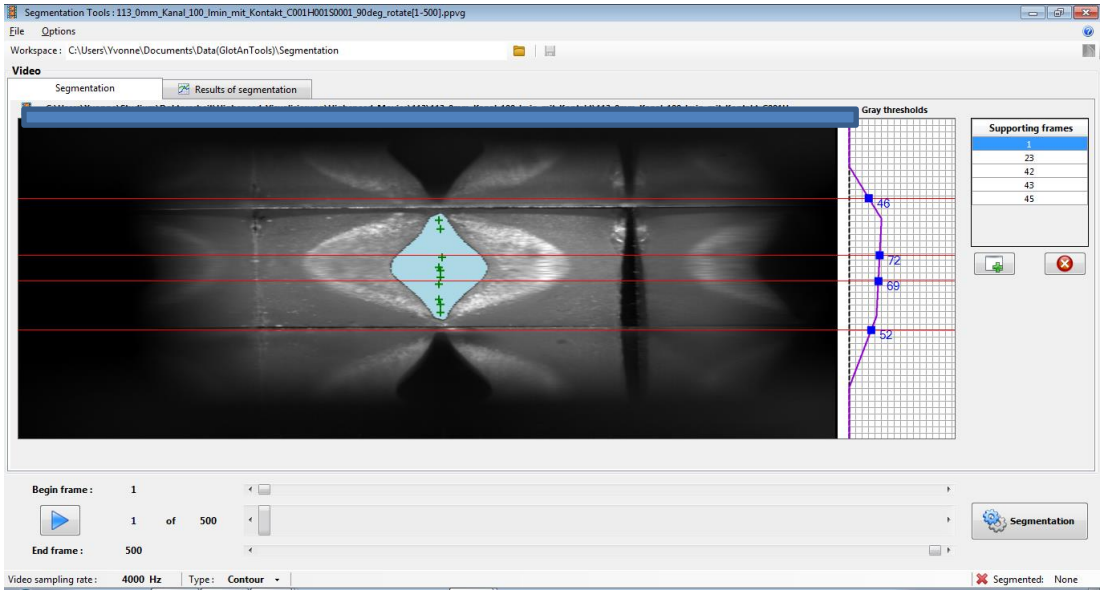


Figure 13 - GAT-Segmentation software with an image of 113 single-layer model 18mm-channel configuration in M2.

Afterwards the segmented videos are analyzed by GAT with the following adjustments:

- Framerate of the recordings in fps
- The expected frequency F0 ranges between 10 and 650 Hz
- Detecting contour
- Sampling rate of 4001 Hz
- Start of the cycle analysis from the beginning, analysis 16 cycles out of 16
- Grey scale threshold detected by the red line (see Figure 13): region of interest in longitudinal direction

[5] Glottis Analysis Tool – Dept. of Otorhinolaryngology, Div. of Phoniatrics & Pediatric Audiology, Waldstraße 1, 91054 Erlangen, Germany

Figure 14 displays the GAW of the 113 model in the 18mm-channel configuration oscillating in mode M2. Seventeen full cycles are detected in the high-speed video.

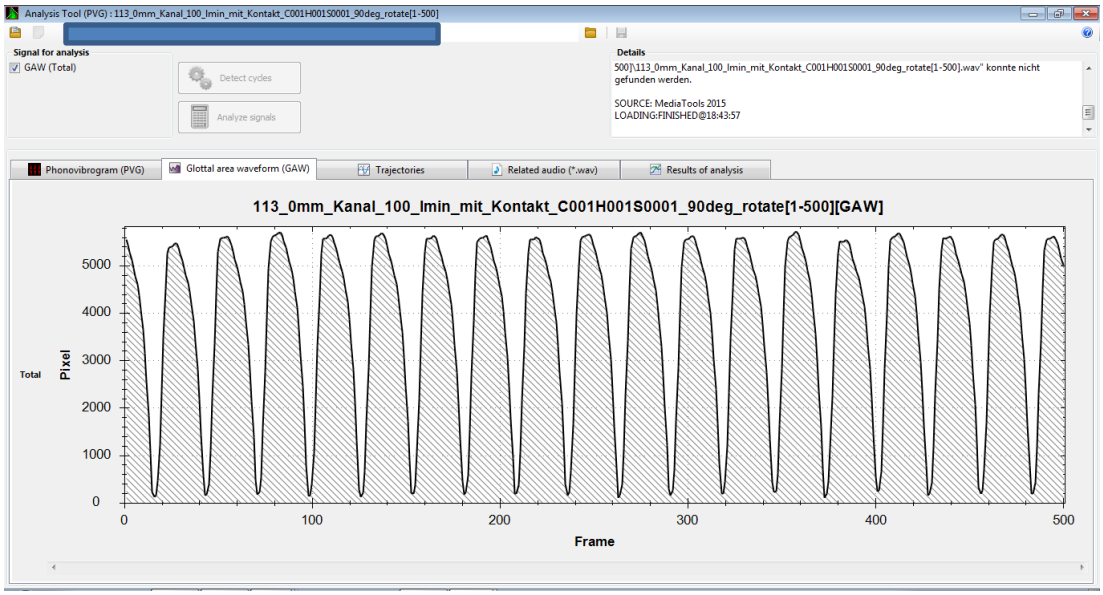


Figure 14 - GAT- Analysis software with an image of 113 single-layer model 18mm-channel configuration in M2. The graph shows the GAW area in pixel during the time in frames.

On the base of the GAW, the program calculates typical objective parameters to evaluate the vocal fold movement. The resulting data is summarized in Table 4 which shows the parameter summary representing the fundamental frequency exemplarily. It is divided in the different models (first column) and the different channel configurations. The third column shows whether the vocal folds have contact (“mit”) or not (“ohne”). The fourth column indicates the channel diameter and the last column displays the values of the parameter (here fundamental frequency).

	Id			Fundamental-Freq(Hz)[GAW][MEAN]
c115	C115_0mm_Kanal_FVF_70_lmin_mit_Kontakt_C001H001S0001_90deg_rotate[1-500]	mit	0	166,739
c115	C115_0mm_Kanal_80_lmin_mit_Kontakt_C001H001S0001_90deg_rotate[1-500]	mit	0	171,42
c114	C114_0mm_Kanal_FVF_85_lmin_mit_Kontakt_C001H001S0001_90deg_rotate[1-500]	mit	0	222,584
c114	C114_0mm_Kanal_90_lmin_mit_Kontakt_C001H001S0001_90deg_rotate[2-500]	mit	0	224,699
114	114_0mm_Kanal_FVF_75_lmin_mit_Kontakt_C001H001S0001_90deg_rotate[1-500]	mit	0	114,342
114	114_0mm_Kanal_105_lmin_mit_Kontakt_C001H001S0001_90deg_rotate[1-500]	mit	0	111,627
113	113_0mm_Kanal_FVF_75_lmin_mit_Kontakt_C001H001S0001_90deg_rotate[1-500]	mit	0	147,936
113	113_0mm_Kanal_100_lmin_mit_Kontakt_C001H001S0001_90deg_rotate[1-500]	mit	0	145,695
c115	C115_5mm_Kanal_80_lmin_mit_Kontakt_C001H001S0001_90deg_rotate[1-500]	mit	5	170,85
c114	C114_5mm_Kanal_100_lmin_mit_Kontakt_C001H001S0001_90deg_rotate[1-500]	mit	5	231,15
113	113_5mm_Kanal_110_lmin_mit_Kontakt_C001H001S0001_90deg_rotate[1-500]	mit	5	143,246
c115	C115_10mm_Kanal_90_lmin_mit_Kontakt_C001H001S0001_90deg_rotate[1-500]	mit	10	173,957
c114	C114_10mm_Kanal_115_lmin_mit_Kontakt_C001H001S0001_90deg_rotate[1-500]	mit	10	231,325
114	114_10mm_Kanal_115_lmin_mit_Kontakt_C001H001S0001_90deg_rotate[1-500]	mit	10	110,908
113	113_10mm_Kanal_115_lmin_mit_Kontakt_C001H001S0001_90deg_rotate[1-500]	mit	10	142,356

Table 4 – Extract from the table of the analyzed parameter of GAT. The columns from left to right: name of the model, description of the measurement, contact of the vocal folds, measured parameter.

In chapter 3.1, selected parameters are illustrated in diagrams for a better overall view with the assistance of data analysis and graphing software Origin 2017®[6].

2.4.2 Parameter choice with description

For an objective and quantitative evaluation on the movement of the vocal folds the following parameters are selected. Every parameter evaluated here is described shortly in the following.

Fundamental frequency

Fundamental frequency is the oscillation frequency of the vocal folds measured in Hz.

Shimmer in %

The parameter shimmer describes the cyclic deviation of the maximum amplitude of GAW at the fundamental frequency in percent.

Mean-jitter in ms

Mean-jitter is the deviation of the cycle duration. The mean-jitter of absolutely periodical vocal fold oscillations is zero. The numeric value can be positive or negative, depending on the direction of the deviation.

Jitter in %

The value of the mean-jitter divided by the average of the cycle duration results in the jitter in percent.

Harmonics to noise ratio

Harmonics to noise ratio describes the ratio of the harmonic amplitudes compared to the noise amplitudes. The value is defined logarithmically and a higher value represents a clearer sound.

Cepstral peak prominence

“CPP (dB) is defined as the difference in amplitude between the cepstral peak and the corresponding value on the regression line computed between 1 millisecond and the maximum quefrency. (i.e., the predicted cepstral magnitude for the quefrency at the cepstral peak).” [46, p. 17]

A high CPP value refers to a high number of harmonics of the fundamental frequency.

[6] Origin 2017 by OriginLab Corporation see https://www.originlab.com/

In this work the CPP is calculated by the formula introduced by Hillenbrand.[7]

MADR

"Maximum-Area-Declination-Rate (MADR) is defined as the maximum amplitude of the negative peak of the 1st derivative of considered signal (here: glottal area waveform or glottal trajectories)." [46, p. 21]

This parameter is a measure of the strength of the produced sound and therefore its loudness or softness [47].

2.4.3 Physiological parameters

Table 5 shows the physiological ranges and the corresponding references of the selected parameters. These are measured and reported by physicians and scientists in healthy patients under the same circumstances as in daily clinical routine [48] [49] [50] [51].

	Physiological measurement	**Source**
Fundamental Frequency female *in Hz*	195 – 375	Ahmad, Yan and Bless (2012) [49]
Fundamental Frequency male *in Hz*	129-259	Warhurst et al. (2014) [50]
Shimmer female *in percent*	0.3-0.28	Schützenberger et al. (2016) [48]
Mean-jitter female *in ms*	0.39-0.34	Schützenberger et al. (2016) [48]
Jitter female *in percent*	6.9-6.1	Schützenberger et al. (2016) [48]
Harmonics to Noise Ratio female *in dB*	14.0-14.5	Schützenberger et al. (2016) [48]
Cepstral Peak Prominence female *in dB*	14,58- 19,82	Monnappa and Balasubramaniam (2015) [52]
Cepstral Peak Prominence male *in dB*	17,71- 21,49	Monnappa and Balasubramaniam (2015) [52]
Maximum Area Declination Rate female	-1096	Schützenberger et al. (2016) [48]
Maximum Area Declination Rate male	(-540) - (-850)	Patel, Dubrovskiy, and Döllinger (2013) [53]

Table 5 - Physiological ranges of GAW based voice parameters and the corresponding references.

[7] For a detailed explanation of the benefit of the usage of the formula of Hillenbrand in case of the calculation of the cepstral peek prominence see Yolanda D. Heman-Ackah (2014).

3 Results and discussion

In the following chapters, the values of the selected parameter are presented. The mean values and the ranges of the parameters are illustrated in a table in each chapter. Also, diagrams give a better overview and enable a better comparison of the parameter subject to the different vocal fold models, their supraglottal channel configuration and their oscillation mode. To enable a comparison to the physiological values of the parameter these values are compared in a bar chart at the end of each chapter.

3.1 Description of mean results

Parameter	**Mean**							
	M1				M2			
	113	114	C114	C115	113	114	C114	C115
Fundamental frequency in dB	135,53	106,14	171,19	137,62	142,02	110,22	219,62	168,40
Shimmer in percent	0,15	0,09	0,70	0,39	0,12	0,13	0,52	0,43
Mean-jitter in ms	0,16	0,16	0,15	0,19	0,18	0,29	0,15	0,13
Jitter in percent	2,24	1,66	2,53	2,68	2,60	3,22	3,34	2,20
Harmonics-to noise ratio in dB	22,28	24,62	17,11	19,33	21,34	22,17	18,28	21,10
Cepstral peak prominence in dB	0,55	0,52	0,64	0,65	1,06	0,66	0,74	0,86
Maximum area declination rate	-257,89	-312,86	-151,81	-122,25	-784,63	-629,01	-780,56	-490,35

Table 6 - Mean values of the parameters of every model in M1 and M2.

Table 6 displays the mean values of all parameters averaged over the supraglottal channel configurations for the four vocal fold model types and M1 and M2. The mean values of the synthetic vocal folds only match the physiological fundamental frequency of the male phonation except for the model 114. Also, the multi-layer model C114 in the M2 mode oscillates with a fundamental frequency characteristic for female phonation. All other parameters show a slight difference in their mean values.

3.2 Fundamental frequency

The next two tables (Table 7and Table 8) show the mean and the range values of all supraglottal channel configurations without FVF for the fundamental frequency in Hz. The tables are subdivided in the M1 and M2, as well as the different vocal fold models.

parameter	mean (all configurations without FVF)							
	M1				M2			
	113	114	C114	C115	113	114	C114	C115
fundamental frequency in dB	135	106	172	137	141	110	212	169

Table 7 - Mean – fundamental frequency in dB of all channel configurations without FVF in M1 and M2.

In Table 7 the mean values of the fundamental frequency in Hz show higher values within the multi-layer models than for the single-layer model in M1 and M2. All models show greater values in M2 than in M1.

parameter	range (all configurations without FVF)							
	M1				M2			
	113	114	C114	C115	113	114	C114	C115
fundamental frequency in dB	10,5	8,0	38,2	15,4	9,7	5,2	73,0	13,5

Table 8 - Range (difference between the maximum value and the minimum value) – fundamental frequency in dB of all channel configurations without FVF in M1 and M2.

As displayed in Table 8 the ranges of the fundamental frequency in Hz of the multi-layer models are greater than the range of the single-layer models in M1 and M2. The influence of the supraglottal channel is therefore greater on the multi-layer models than on the single-layer models. In M2 the range of the single-layer models is smaller than in M1. This demonstrates the stabilizing effect of M2 in comparison to M1 on the oscillation of the vocal folds. C114 show an increase of the range in M2 in comparison to M1, which is contrariwise in C115.

In the following, Figure 15 shows the fundamental frequency in Hz for a better visual analysis of the vocal fold model movements. The different models C115, C114, 114 and 113 can be distinguished by different colors and signs of the single graphs.

In Figure 15, the x-axis shows the supraglottal channel configurations from 18mm up to 78mm, no channel and the supraglottal channel with FVFs. The y-axis shows the fundamental frequency in Hz.

All models and channel configurations are recorded for two different oscillation modes. M1 (left side of the diagram) is without, while M2 (right side of the diagram) is with contact of the vocal folds in each oscillation cycle.

The data points represent the mean over all cycles of the detected high-speed recordings. The standard deviation is displayed in the whiskers of the boxplot-like display.

The solid lines show the trend line of the respective model, while the dotted lines represent linear devolution between the single channel configurations. The channel configuration 8 with FVFs is excluded from the calculation of the regression line.

In M2, data is missing for the multi-layer models in the channel configuration 8. This is because M2 did not occur in this channel configuration with the multi-layer models. The data for M1 with model C114 and with FVF are missing as well [5, p. 61].

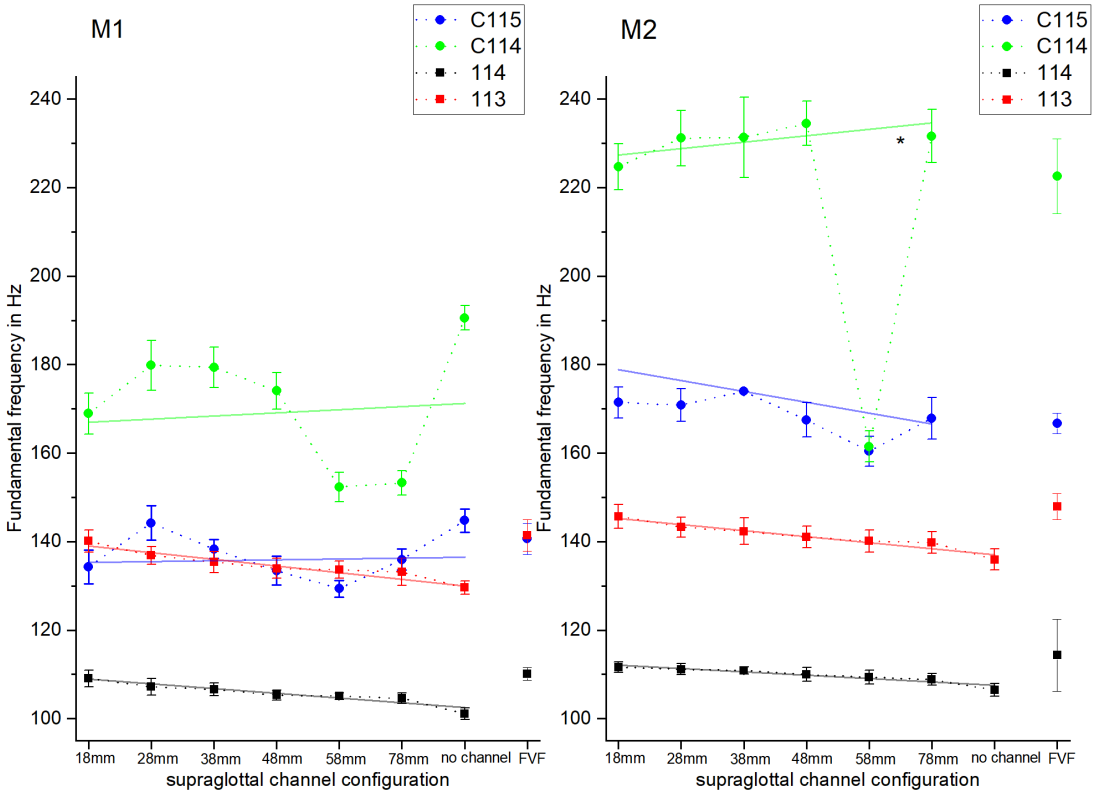


Figure 15 – Mean fundamental frequency and standard deviation [in Hz] in M1 (left diagram) and M2 (right diagram) for different supraglottal channel configurations.

The C115 and C114 models have higher values in M1 and M2 than the single-layer models. The standard deviation of the single-layer models is smaller and it is almost constant in all supraglottal channel configurations and both oscillation modes. The greater standard deviation of the fundamental frequency of the multi-layer models indicates a correlation between the inner structure of the models and the oscillation mode. M2 exhibits a greater standard deviation in every model type in comparison to M1. The increase of the standard deviation from M1 to M2 in the multi-layer models is larger than the one of the single-layer models. This may result from of the stabilization of the single-layer model trough a better contact of the vocal folds during the oscillation in M2 (see Figure 16). The composition of the synthetic vocal folds affects the stability of the fundamental frequency.

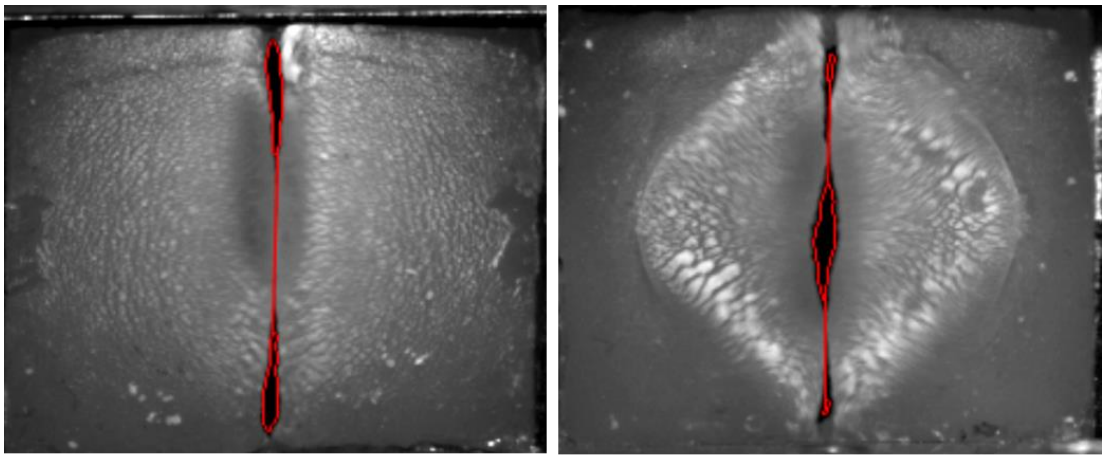


Figure 16 - Contact of the vocal folds during the glottis closure – Single (left side) and multi-layer models (right side) in direct comparison. Pictures according to [5].

In M2, the increase of the fundamental frequency of the multi-layers is higher compared with M1, while the single-layer models show a more moderate increase. Also, this could be related to the impact of the structure of the multi-layer models. Their softer cover layer in combination with the stiff ligament leads to an increase in their frequency.

Among the multi-layer models, C114 shows higher fundamental frequencies than C115. Furthermore, C115 has more constant values, but the change of the oscillation mode produces a larger increase in frequency for C114. The standard deviation for the C114 model is greater than the one of the C115 model for both modes as well. This is discussed more detailed later in chapter 3.4 and 3.5 within the jitter in percent and the mean-jitter.

C114 has an outlier in the *38mm* channel configuration of M2, which is not considered in the generated trend line.

The 113 model has a higher fundamental frequency than the 114 model, but both single-layer models follow the same decreasing trend.

For the single layer models a slight decrease in the frequency for the increasing supraglottal channel diameter in both oscillation modes is measured. The multi-layer models show this behave only in M2 in C115.

The single-layer models show a decrease in the *no channel* channel configuration in both modes compared to the 1*8mm* channel configuration, which displays the context to the flow properties of the supraglottal channel.

Adding the FVF, no significant change in the frequency compared to the 1*8mm* channel configuration can be detected.

In the following, Figure 17 shows the results of the synthetic vocal folds of the fundamental frequency in Hz the realistic reproduction of physiological acoustics in comparison to the oscillations of human vocal folds. This is based on clinically relevant and objective parameters of the high-speed videos of the synthetic vocal fold movement.

Subsequently, the data by Ahmad et al., Warhust et al. and the data of the synthetic larynx models C115, C114, 114 and 113 will be compared regarding to each parameter [5] [49] [50]. Additionally, the different gender (female [49] and male [50]) are subdivided to gain more information about the behave of the synthetic vocal fold models.

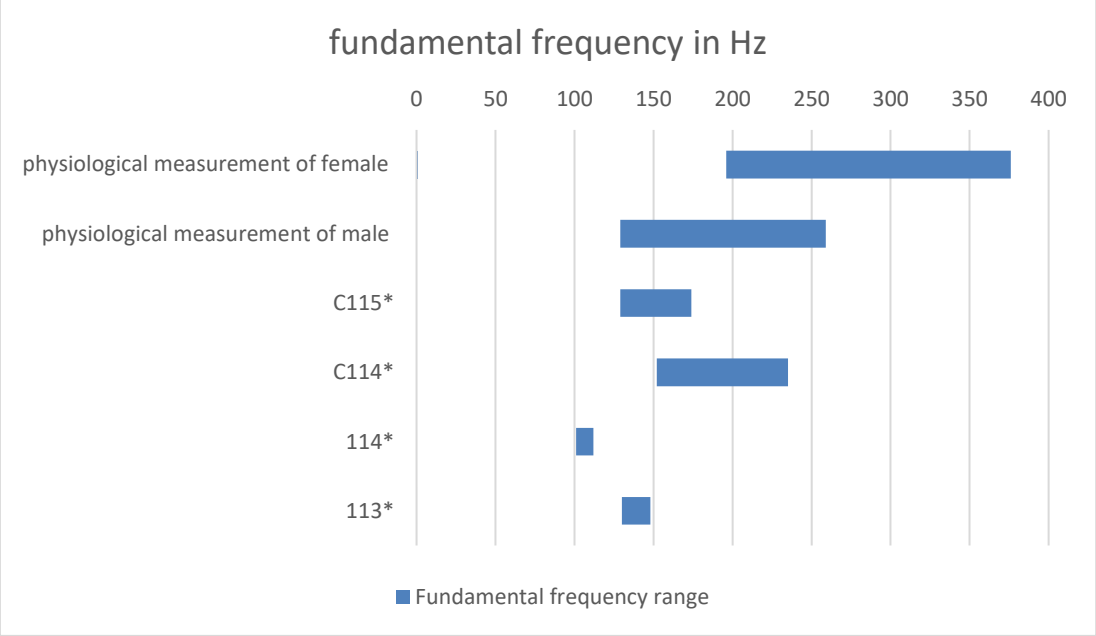


Figure 17 - Synthetic vocal fold model data in comparison to the physiological human data: Ranges of the fundamental frequency in Hz (Physiological female and male versus 113, 114, C114 and C115 model) [5] [49] [50].

Figure 17 shows the range of the fundamental frequency parameter of the different synthetic vocal fold model types in comparison with the physiological values of male [50] and female [49] subjects.

As illustrated, the values of the fundamental frequency for the male phonation and the multi-layer as well as the 113 models overlap. The values of the 114 models are slightly lower than the rest.

Only the C114 model shows an overlap with the physiological female group. The majority of the single- and multi-layer models match with the physiological ranges of the male subjects. The ranges for physiological phonation are larger than the range of the synthetic models.

The multi-layer models oscillate in a similar fundamental frequency to the male vocal folds. This is to say that the single-layer model 114 does not replicate the physiological voice production as closely as the other synthetic models.

3.3 Shimmer in percent

The next two tables (Table 9 and Table 10) show the mean and the range values of all supraglottal channel configurations without FVF for the shimmer in percent. The tables are subdivided in the M1 and M2, as well as the different vocal fold models.

parameter	**mean (all configurations without FVF)**							
	M1				M2			
	113	114	C114	C115	113	114	C114	C115
shimmer in percent	0,13	0,08	0,62	0,33	0,10	0,05	0,46	0,47

Table 9 - Mean – shimmer in percent of all channel configurations without FVF in M1 and M2.

In Table 9 the mean values of the shimmer in percent show higher values for the multi-layer models than for the single-layer model in M1 and M2. All single-layer models show smaller values in M2 than in M1. C114 has smaller mean values in M2 than in M1, while C115 behaves the other way around.

Parameter	**range (all configurations without FVF)**							
	M1				M2			
	113	114	C114	C115	113	114	C114	C115
shimmer in percent	0,15	0,07	1,72	0,70	0,16	0,04	0,56	0,36

Table 10 - Range (difference between the maximum value and the minimum value) – shimmer in percent of all channel configurations without FVF in M1 and M2.

The ranges of the shimmer in percent, displayed in Table 10 ,of the multi-layer models are greater than the ranges of the single-layer models in M1 and M2. The single-layer models show only a slight change of the range comparing M1 to M2, while the multi-layer models in M2 divide the range values of M1 nearly in half. This demonstrates the stabilizing effect of M2 for the multi-layer models.

In the following, Figure 18 illustrates the shimmer in percent for a better visual analysis of the vocal fold model movements. The different models C115, C114, 114 and 113 can be distinguished by different colors and signs of the single graphs.

In Figure 18 the x-axis shows the supraglottal channel configurations from 18mm up to 78mm, no channel and the supraglottal channel with FVFs. The y-axis shows the shimmer in percent.

All models and channel configurations are recorded for two different oscillation modes. M1 (left side of the diagram) is without, while M2 (right side of the diagram) is with contact between the vocal folds in each oscillation cycle.

The data points represent the mean over all cycles of the detected high-speed recordings.

The solid lines show the trend line of the respective model, while the dotted lines represent linear devolution between the single channel configurations. The channel configuration 8 with FVFs is excluded from the calculation of the regression line.

In M2, data is missing for the multi-layer models in the channel configuration 8. This is because M2 did not occur in this channel configuration with multi-layer models. The data for M1 with model C114 and with FVF are missing as well [5, p. 61].

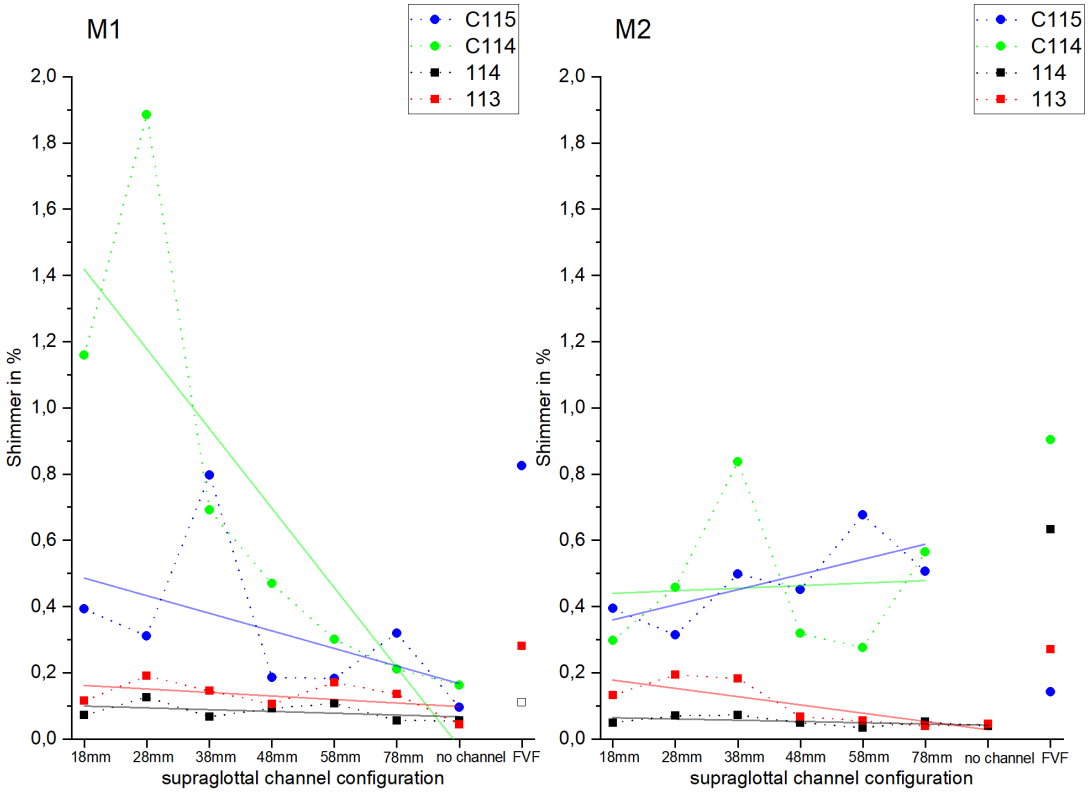


Figure 18 - Shimmer [in %] in M1 (left diagram) and M2 (right diagram) for different supraglottal channel configurations.

The values of the single-layer models are lower than these of the multi-layer models. Also, the devolution of the Shimmer for the single-layer models for the increasing channel diameter is smoother than those of the multi-layer models. The range for multi-layer models is larger than the range for the single layer models.

Considering all models, M2 has a smaller range of values (Δ: 0,87) than M1 (absolute value: 1,84).[8]

[8] M1: highest value 0,904% (C114 in channel configuration 2), lowest value 0,034% (113 in channel configuration 7)

C114 has higher shimmer values especially in M1 in comparison to C115.

Considering the single-layer models, 113 has higher Shimmer values than 114 in M1 and M2. The first three channel configurations in M1 and M2 show increasing values in 113. In M2 the *48mm, 58mm, 78mm and no channel* channel configurations show steady values. Within smaller channel diameter the flow region of the supraglottal area develops a higher shimmer, which indicates a larger impact of flow fluctuations on the vocal fold motion and therefore leads to higher perturbations of the amplitude. Thus, the Shimmer decreases with increasing channel diameter. This relation is demonstrated for all model types by the fact that the Shimmer of the *no channel* channel configuration exhibits the smallest values in comparison to the other channel configurations. Therefore, the perturbation of the amplitude becomes smaller. It can be associated with a less rough sound signal without a supraglottal channel in comparison to a supraglottal channel with small diameter.

The multi-layer and single layer models in M1 exhibit a decreasing trend in general. In M2 the single-layer models show a decreasing trend, whereas the multi-layer models in M2show increasing trend.

The *no channel* channel configuration has lower values in all model types than the *18mm* channel configuration. The influence of the supraglottal channel in M1 is greater on the multi-layer models, because the difference between the *no channel* channel configuration and the *18mm* channel configuration in C114 and C115 is significant. The removal of the supraglottal channel leads to a smaller perturbation. The supraglottal channel seems to have a greater influence on the flow field regarding the multi-layer than the single-layer models.

The values of the *FVF* channel configuration differ from the *18mm* channel configuration in M1 and M2. In M1 the shimmer with the FVF is higher in every model type in comparison to the *18mm* channel configuration without a FVF. The same applies to M2. This means that the *FVF* channel configuration produces more roughness in the acoustic signal [54].

In the following, Figure 19 illustrates the results of the synthetic vocal folds of the shimmer in percent and the realistic reproduction of physiological acoustics in comparison to the oscillations of human vocal folds. This is based on clinically relevant and objective parameters of the high-speed videos of the synthetic vocal fold movement.

M2: highest value 1,886% (C114 in channel configuration 3), lowest value 0,044% (113 in channel configuration 7)

Subsequently, the data by Schützenberger et al. and the data of the synthetic larynx models C115, C114, 114 and 113 will be compared regarding to each parameter [5] [48].

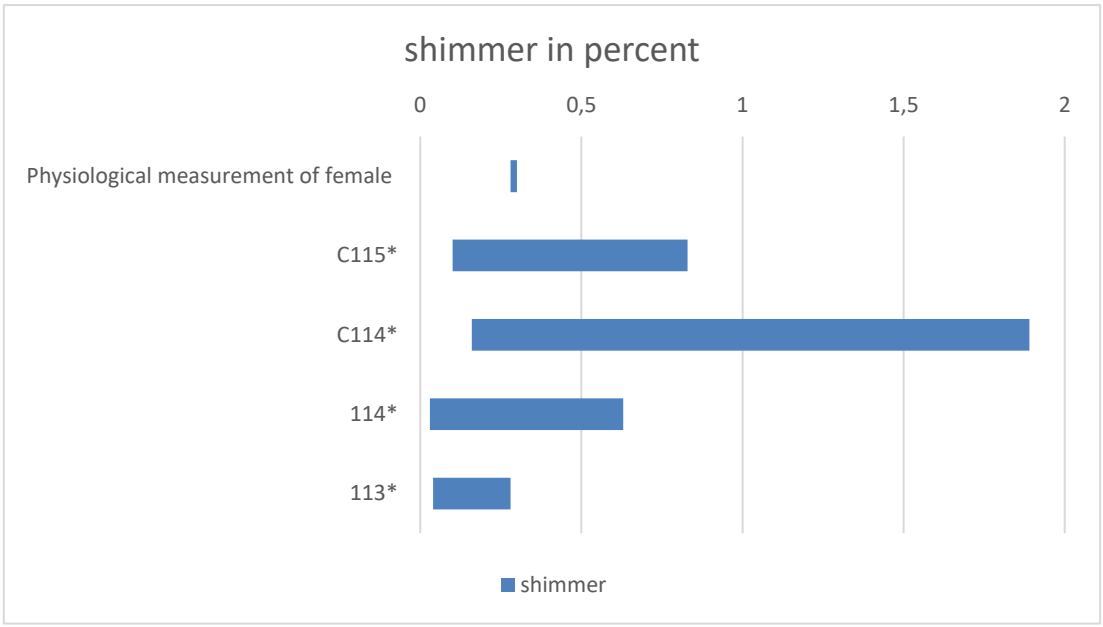


Figure 19 - Synthetic vocal fold model data in comparison to the physiological human data: Ranges of the shimmer in percent (Physiological female and male versus 113, 114, C114 and C115 model) [5] [48].

The synthetic vocal folds exhibit larger ranges of the shimmer than the physiological vocal folds. The ranges of the synthetic models cover the physiological range (except for 113).

3.4 Mean-jitter

The next two tables (Table 11 and Table 12) show the mean and the range values of all supraglottal channel configurations without FVF for the mean-jitter in ms. The tables are subdivided in the M1 and M2, as well as the different vocal fold models.

parameter	**mean (all configurations without FVF)**							
	M1				M2			
	113	114	C114	C115	113	114	C114	C115
mean-jitter in ms	0,15	0,16	0,15	0,19	0,19	0,14	0,14	0,15

Table 11 - Mean – mean-jitter in ms of all channel configurations without FVF in M1 and M2.

In Table 11 the mean values of the mean-jitter in ms in M1 show higher values than in M2 in all model types, but all in all the aberrance in between the models is small.

parameter	range (all configurations without FVF)							
	M1				M2			
	113	114	C114	C115	113	114	C114	C115
mean-jitter in ms	0,15	0,23	0,18	0,20	0,13	0,19	0,17	0,22

Table 12 - Range (difference between the maximum value and the minimum value) – mean-jitter in percent of all channel configurations without FVF in M1 and M2.

The ranges of the mean-jitter in ms are, displayed in Table 12, of all model types excluding C115 are greater in M1 than in M2. The influence of the supraglottal channel is therefore greater in M1 than in M2 and therefore is the oscillation in M2 more stable than in M1.

The softer models (114 and C115) show a wider range than the harder ones (113 and C114) in M1 and M2. This result from the fact that the softer tissue reacts more sensitive to small asymmetries of the flow, which leads to a temporary change of the length of the cycle.

In the following, Figure 20 illustrates the mean-jitter in ms for a better visual analysis of the vocal fold model movements. The different models C115, C114, 114 and 113 can be distinguished by different colors and signs of the single graphs.

In Figure 20, the x-axis shows the supraglottal channel configurations from 18mm up to 78mm, no channel and the supraglottal channel with FVFs. The y-axis shows the mean-jitter in ms.

All models and channel configurations are recorded for two different oscillation modes. M1 (left side of the diagram) is without, while M2 (right side of the diagram) is with contact between the vocal folds in each oscillation cycle.

The data points represent the mean over all cycles of the detected high-speed recordings.

The solid lines show the trend line of the respective model, while the dotted lines represent linear devolution between the single channel configurations. The channel configuration 8 with FVFs is excluded from the calculation of the regression line.

In M2, data is missing for the multi-layer models in the channel configuration 8. This is because M2 did not occur in this channel configuration with multi-layer models. The data for M1 with model C114 and with FVF are missing as well [5, p. 61].

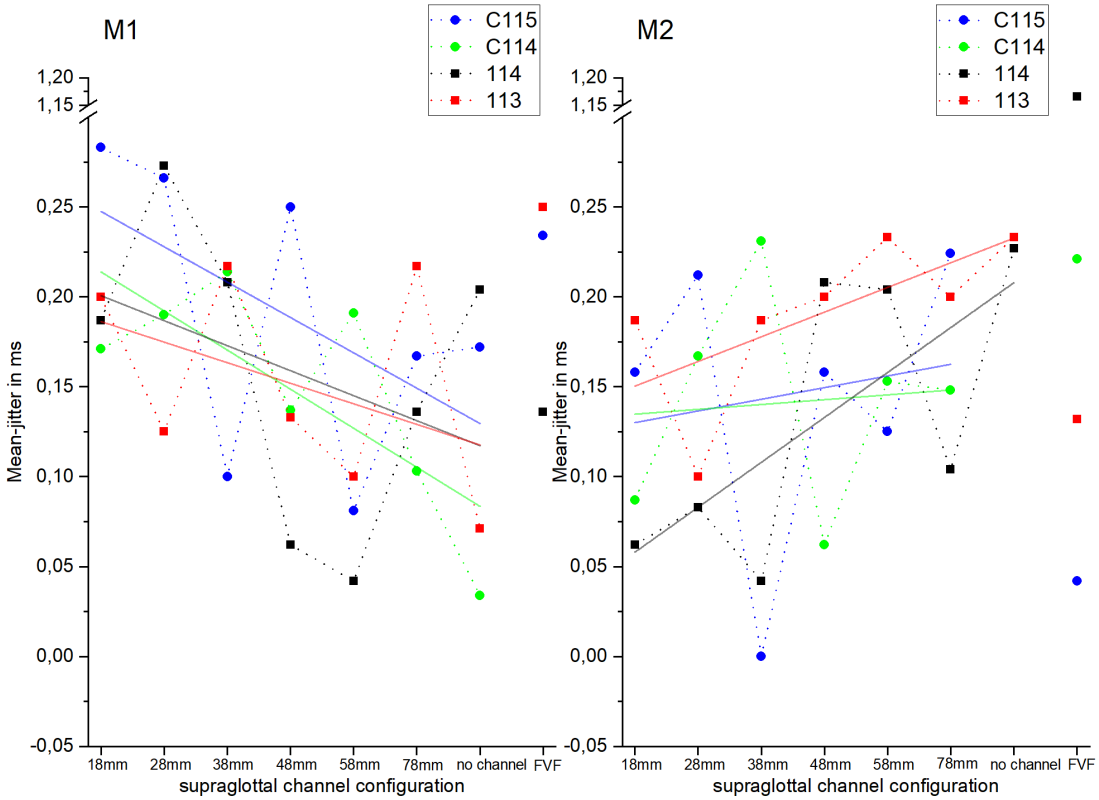


Figure 20 - Mean-jitter [in ms] in M1 (left diagram) and M2 (right diagram) for different supraglottal channel configurations.

Only small differences can be observed between all vocal fold models.

The trends of the mean-jitter as function of the lateral diameter of the supraglottal channel have a different slope in M1 and M2. Whereas the slope is positive in M1, the slope is negative in M2 for all vocal fold models. This may result from a smaller fluctuation of the flow for wider lateral channel diameters in M1. This leads to a reduced variability in the length of the cycle. In M2 this correlation is inverted.

The C114, C115 and 113 models have lower values in the *no channel* channel configuration than in the *18mm* channel configuration in M1. As a conclusion, without the supraglottal channel the perturbation of the period length is lower which can have a positive effect on the tonal sound generation. The opposite is shown in M2, where the *no channel* channel configuration has higher values than the *18mm* channel configuration. The oscillation modes seem to have a great influence on the mean-jitter within the single-layer models.

In the following, Figure 21 illustrates the results of the synthetic vocal folds of the mean-jitter in ms and the realistic reproduction of physiological acoustics in comparison to the oscillations of human vocal folds. This is based on clinically relevant and objective parameters of the high-speed videos of the synthetic vocal fold movement.

Subsequently, the data by Schützenberger et al. and the data of the synthetic larynx models C115, C114, 114 and 113 will be compared regarding to each parameter [48] [5].

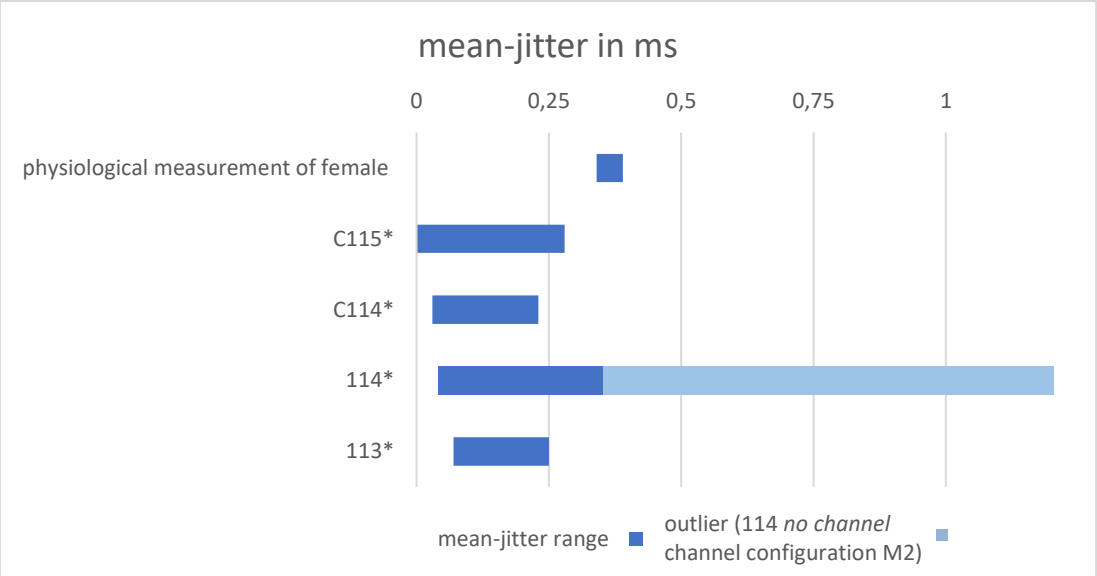


Figure 21 - Synthetic vocal fold model data in comparison to the physiological human data: Ranges of the mean-jitter in ms (Physiological female and male versus 113, 114, C114 and C115 model) [48] [5].

The physiological mean-jitter values are about twice as high as the values of the models. Which means a more constant oscillation of the synthetic vocal folds in comparison to the human vocal folds. 114 has an outlier, which is shown in Figure 21. The total range of the physiological mean-jitter is smaller than the mean-jitter of the single- and multi-layer models. Therefore, the supraglottal channel dimensions have a large impact and thus they effect the aerodynamic perturbations referring the period length of the oscillating vocal folds. An effect of the fundamental frequency on the jitter is reported by Brockmann et al. [55]. An increase of the fundamental frequency leads to a decrease of the jitter because of the stiffening of the vocal folds during the increase of the fundamental frequency [56].

A direct comparison can be made under the same measurement and environmental conditions only. The multi-layer models have a greater fundamental frequency and therefore a smaller jitter than the single-layer models[9] (see chapter 3.2). Hence, the correlation of the fundamental frequency and the jitter is not obvious in 113, because of the very periodical oscillation of this model. The jitter of the female subjects [48] is higher than the ones of the synthetic models, which is caused by a less precise voice quality (no professional singer) in comparison to the very periodical oscillation of the synthetic vocal folds.

[9] outlier in 114 excluded

3.5 Jitter in percent

The next two tables (Table 13 and Table 14) show the mean and the range values of all supraglottal channel configurations without FVF for the jitter in percent. The tables are subdivided in the M1 and M2, as well as the different vocal fold models.

parameter	**mean (all configurations without FVF)**							
	M1				M2			
	113	114	C114	C115	113	114	C114	C115
jitter in percent	2,05	1,68	2,46	2,59	2,70	1,54	3,08	2,46

Table 13 – Mean – jitter in percent of all channel configurations without FVF in M1 and M2.

In Table 13 the mean of the single-layer models shows a greater distance in M2 than in M1, the same correlation takes place within the multi-layer models.

parameter	**range (all configurations without FVF)**							
	M1				M2			
	113	114	C114	C115	113	114	C114	C115
jitter in percent	2,01	2,48	3,19	2,78	1,84	1,96	3,88	3,75

Table 14 - Range (difference between the maximum value and the minimum value) - Jitter in percent of all channel configurations without FVF in M1 and M2.

As displayed in Table 14, the range of the single-layer models is higher in M1 than in M2, while the multi-layer models show a contrary correlation. In general, the range of the multi-layer models is higher than the range of the single-layer models, which demonstrates the greater influence of the supraglottal channel on the multi-layer models than on the single-layer models.

In the following, Figure 22 shows the jitter in percent for a better visual analysis of the vocal fold model movements. The different models C115, C114, 114 and 113 can be distinguished by different colors and signs of the single graphs.

In Figure 22, the x-axis shows the supraglottal channel configurations from 18mm up to 78mm, no channel and the supraglottal channel with FVFs. The y-axis shows the jitter in percent.

All models and channel configurations are recorded for two different oscillation modes. M1 (left side of the diagram) is without, while M2 (right side of the diagram) is with contact between the vocal folds in each oscillation cycle.

The data points represent the mean over all cycles of the detected high-speed recordings.

The solid lines show the trend line of the respective model, while the dotted lines represent linear devolution between the single channel configurations. The channel configuration 8 with FVFs is excluded from the calculation of the regression line.

In M2 data is missing for the multi-layer models in the *no channel* channel configuration. This is because M2 did not occur in this channel configuration with multi-layer models. The data for M1 with model C114 and with FVF are missing as well [5, p. 61].

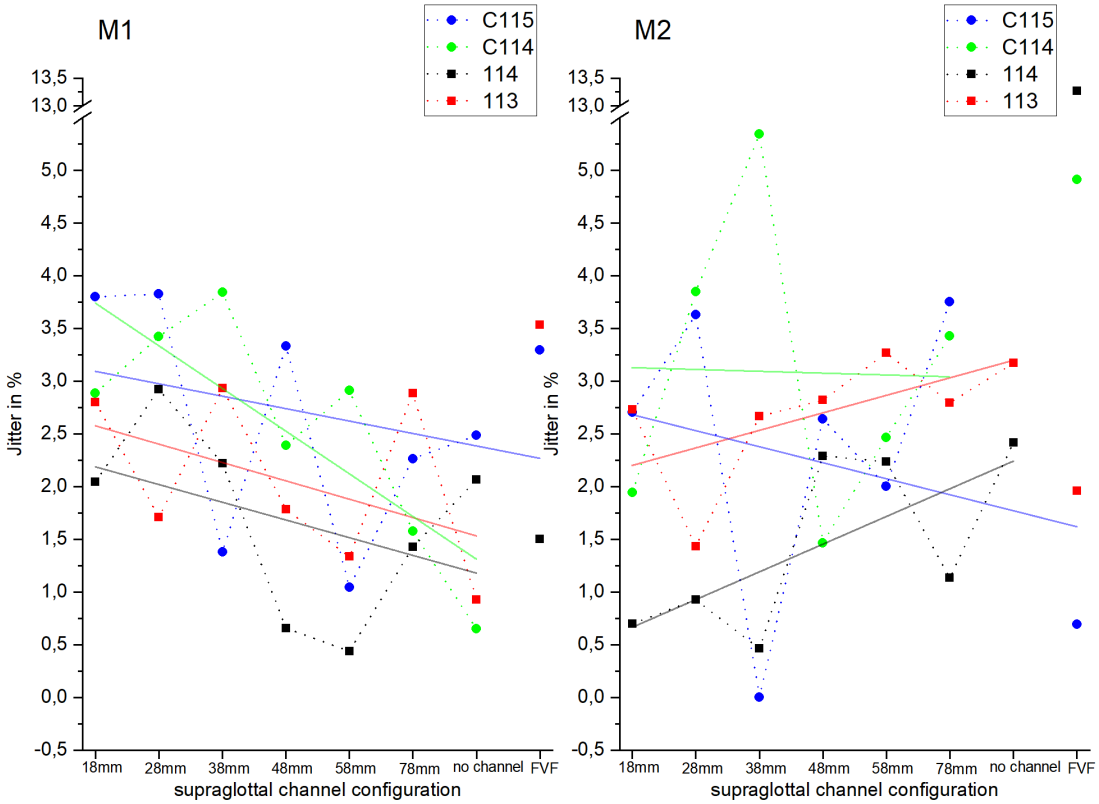


Figure 22 – Jitter [in %] in M1 (left diagram) and M2 (right diagram) for different supraglottal channel configurations.

The results of oscillation modes, the model types and the change of the channel configurations of the jitter in percent can be compared to the results of the mean-jitter (see chapter 3.4).

The global trend shows in M1 a decreasing trend in all model types, while in M2 the trend is increasing only for the single-layer models.

In the following, Figure 23 illustrates the results of the synthetic vocal folds of the jitter in percent and the realistic reproduction of physiological acoustics in comparison to the oscillations of human vocal folds. This is based on clinically relevant and objective parameters of the high-speed videos of the synthetic vocal fold movement.

Subsequently, the data by Schützenberger et al. and the data of the synthetic larynx models C115, C114, 114 and 113 will be compared regarding to each parameter [48] [5].

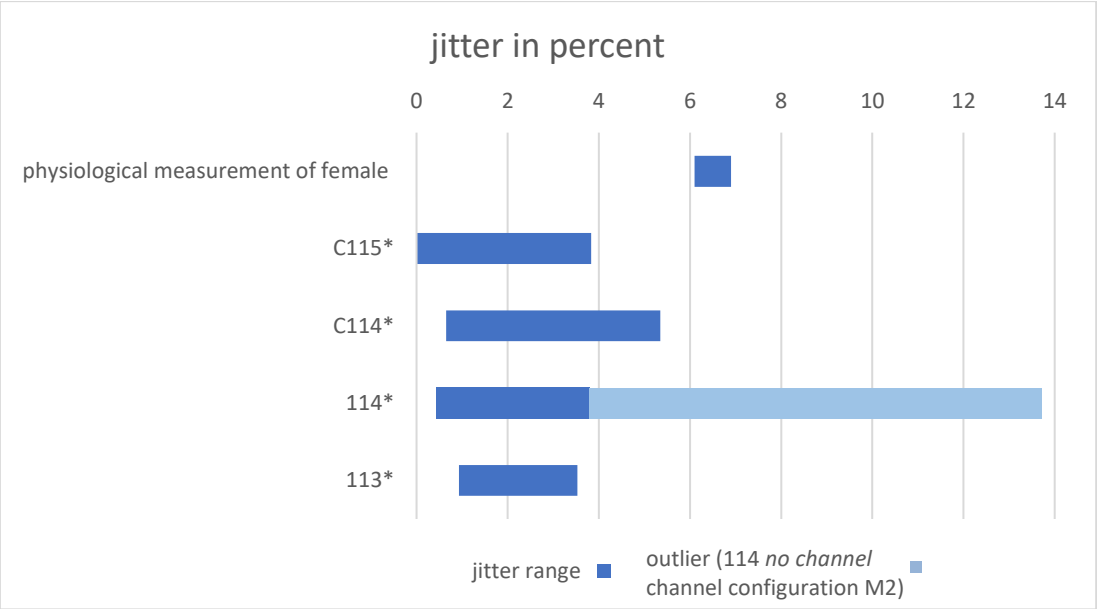


Figure 23 - Synthetic vocal fold model data in comparison to the physiological human data: Ranges of the jitter in percent (Physiological female and male versus 113, 114, C114 and C115 model) [48] [5].

The 114 model has an outlier, which is shown in Figure 21. All values of the models (excluding the outlier in 114) are lower than the range of the physiological measurements. The greater range of 114 could be explained by an outlier in the *FVF* channel configuration. This outlier might result from damage to the synthetic model during the long measurements or a non-exact adjustment of the FVF inside the supraglottal channel.

3.6 Harmonics to noise ratio

The next two tables (Table 15 and Table 16) show the mean and the range values of all supraglottal channel configurations without FVF for the harmonics to noise ratio in dB. The tables are subdivided in the M1 and M2, as well as the different vocal fold models.

parameter	**mean (all configurations without FVF)**							
	M1				M2			
	113	114	C114	C115	113	114	C114	C115
harmonic to noise ratio in dB	22,7	24,9	17,7	19,7	21,5	24,3	18,7	20,9

Table 15 –Mean - Harmonic to noise ratio of all channel configurations without FVF in M1 and M2.

In Table 15 the mean of the single-layer models is higher in M1 than in M2. For the multi-layer models the mean is higher in M2 than in M1. In general, the mean values of the single-layer models are higher than the mean values of the multi-layer models.

parameter	range (all configurations without FVF)							
	M1				M2			
	113	114	C114	C115	113	114	C114	C115
harmonic to noise ratio in dB	2,44	3,77	10,73	4,88	3,52	1,90	3,73	6,70

Table 16 - Range (difference between the maximum value and the minimum value) - Harmonic to noise ratio in dB of all channel configurations without FVF in m1 and M2.

For the multi-layer models the ranges are wider than ranges of the single-layer models in each oscillation mode, as displayed in Table 16. The influence of the supraglottal channel is therefore greater on the multi-layer models than on the single-layer models.

In the following, Figure 24 illustrates the harmonics to noise ratio in dB for a better visual analysis of the vocal fold model movements. The different models C115, C114, 114 and 113 can be distinguished by different colors and signs of the single graphs.

In Figure 24, the x-axis shows the supraglottal channel configurations from 18mm up to 78mm, no channel and the supraglottal channel with FVFs. The y-axis shows the harmonics to noise ratio in dB.

All models and channel configurations are recorded for two different oscillation modes. M1 (left side of the diagram) is without, while M2 (right side of the diagram) is with contact between the vocal folds in each oscillation cycle.

The data points represent the mean over all cycles of the detected high-speed recordings.

The solid lines show the trend line of the respective model, while the dotted lines represent linear devolution between the single channel configurations. The *FVF* channel configuration is excluded from the calculation of the regression line.

In M2, data is missing for the multi-layer models in the *FVF* channel configuration. This is because M2 did not occur in this channel configuration with multi-layer models. The data for M1 with model C114 and with FVF are missing as well [5, p. 61].

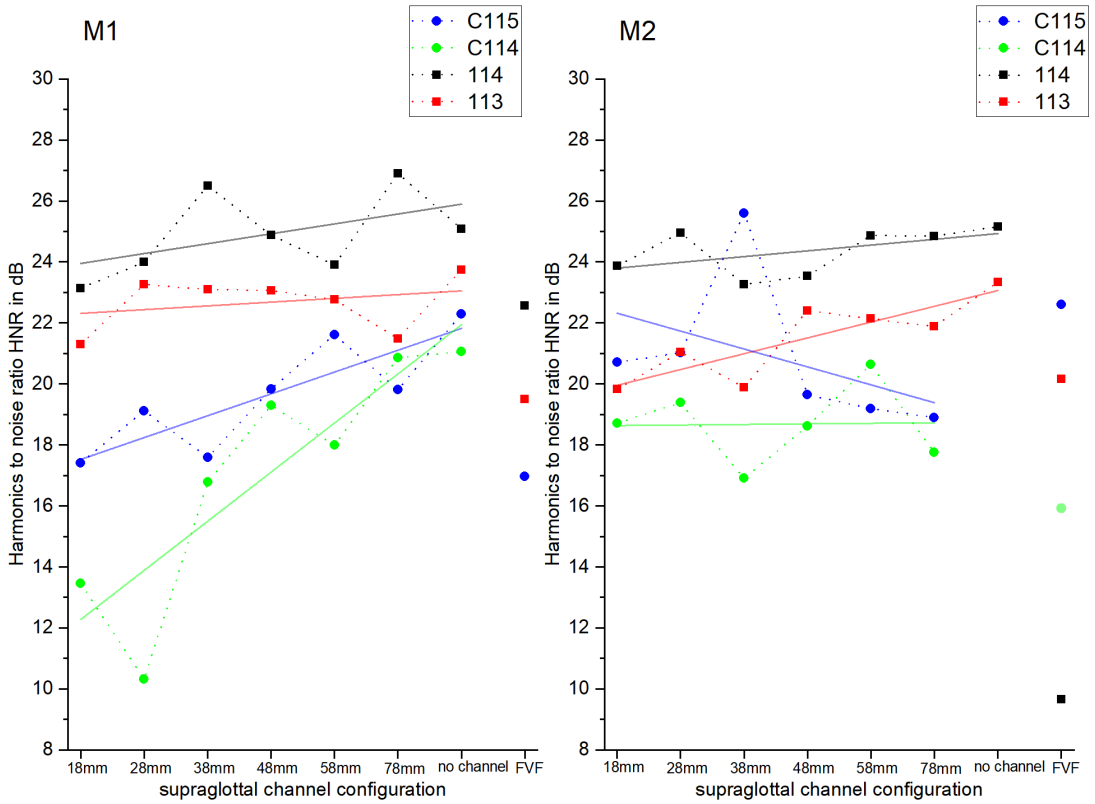


Figure 24 – Harmonics to noise ratio [in dB] in M1 (left diagram) and M2 (right diagram) for different supraglottal channel configurations.

The single-layer models have higher values in M1 than the multi-layer models.

The values of all models increase in M1 and M2 for increasing channel diameter, except the multi-layer models which generate a decreasing (C115) or constant (C114) trend in M2.

The C114 model shows lower values than the C115 model in M1 and M2. They follow the same general trend only in M1.

The values of the single-layer models show no great change over all channel configurations and the values of the 114 model are higher than ones of the 113 model during the change of the supraglottal channel diameter.

Comparing the general trends, the single-layer models show a slightly increasing trend in M1 and M2.

In M1 and M2, the values of the *no channel* channel configuration are higher than the values of the *18mm* channel configuration. Thus, the harmonic content of the GAW increases with increasing channel diameter.

The insertion of the FVFs in M1 shows higher values than in the *18mm* channel configuration. The FVF have a positive effect on the HNR.

In the following, Figure 25 illustrates the results of the synthetic vocal folds of the harmonics to noise ratio in dB and the realistic reproduction of physiological acoustics in comparison to the oscillations of human vocal folds. This is based on clinically relevant and objective parameters of the high-speed videos of the synthetic vocal fold movement.

Subsequently, the data by Schützenberger et al. and the data of the synthetic larynx models C115, C114, 114 and 113 will be compared regarding to each parameter [48] [5].

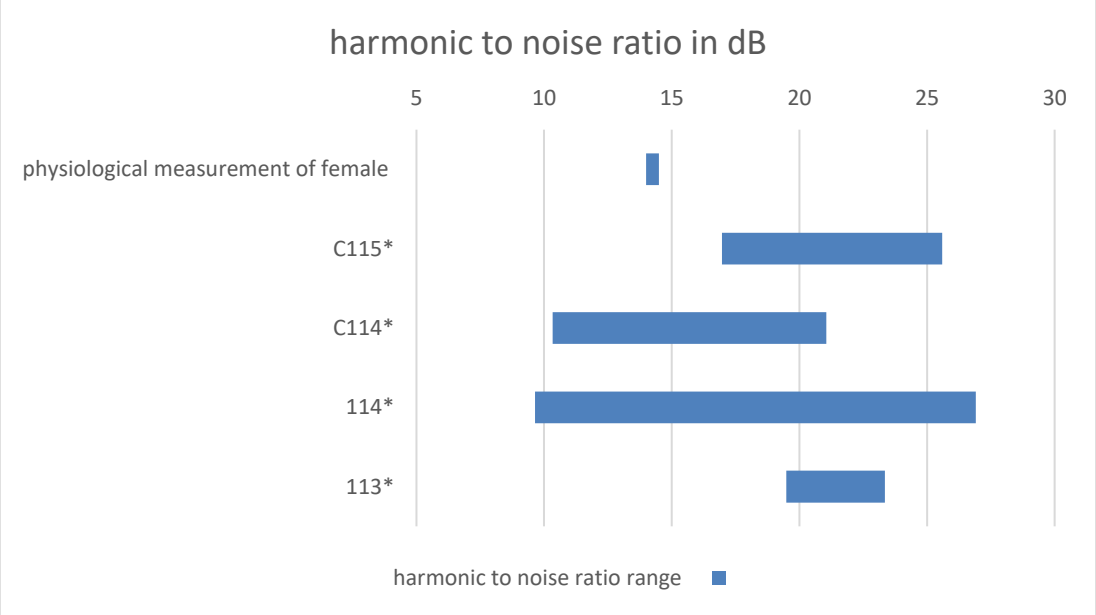


Figure 25 - Synthetic vocal fold model data in comparison to the physiological human data: Ranges of the harmonics to noise ratio in dB (Physiological female and male versus 113, 114, C114 and C115 model) [48] [5].

The values of the physiological and synthetic vocal folds are in the same range. The range of the synthetic models is larger, originating from a large influence of the supraglottal channel.

3.7 Cepstral peak prominence

The next two tables (Table 17 and Table 18) show the mean and the range values of all supraglottal channel configurations without FVF for the cepstral peak prominence in dB. The tables are subdivided in the M1 and M2, as well as the different vocal fold models.

parameter	**mean (all configurations without FVF)**							
	M1				M2			
	113	114	C114	C115	113	114	C114	C115
cepstral peak prominence in dB	0,54	0,51	0,63	0,66	1,05	0,58	0,742	0,83

Table 17 - Mean - cepstral peak prominence of all channel configurations without FVF in M1 and M2.

In Table 17 the mean values in general are higher in M1 than in M2. The mean values of the single-layer models are lower than the ones of the multi-layer models in M1.

The total glottal closure of the vocal folds stabilizes the process of phonation, which is perceivable by the increase of the mean of the CPP in M2. This higher periodicity leads to a larger number of harmonics in the phonation. The multi-layer models have no total closure of the glottal gap and therefore the CPP values in M2 are lower than those of the single-layer models (except 114). This shows the importance of the total glottal closure during the phonation. [57, p. 2205]

parameter	**range (all configurations without FVF)**							
	M1				M2			
	113	114	C114	C115	113	114	C114	C115
cepstral peak prominence in dB	0,201	0,402	0,415	0,654	1,182	1,017	0,731	0,727

Table 18 - Range (difference between the maximum value and the minimum value) – cepstral peak prominence in dB of all channel configurations without FVF in M1 and M2.

The range and the mean of the values of all models in M2 are larger than in M1, see Table 17 and Table 18. The single-layer models have a greater range of values than the multi-layer models in M2.

In the following, Figure 26 illustrates the cepstral peak prominence in dB for a better visual analysis of the vocal fold model movements. The different models C115, C114, 114 and 113 can be distinguished by different colors and signs of the single graphs.

In Figure 26, the x-axis shows the supraglottal channel configurations from 18mm up to 78mm, no channel and the supraglottal channel with FVFs. The y-axis shows the cepstral peak prominence.

All models and channel configurations are recorded for two different oscillation modes. M1 (left side of the diagram) is without, while M2 (right side of the diagram) is with contact between the vocal folds in each oscillation cycle.

The data points represent the mean overall cycles of the detected high-speed recordings.

The solid lines show the trend line of the respective model, while the dotted lines represent linear devolution between the single channel configurations. The *FVF* channel configuration is excluded from the calculation of the regression line.

In M2, there is data missing for the multi-layer models in the *FVF* channel configuration. This is because M2 did not occur in this channel configuration with multi-layer models. The data for M1 with model C114 and with FVF are missing as well [5, p. 61].

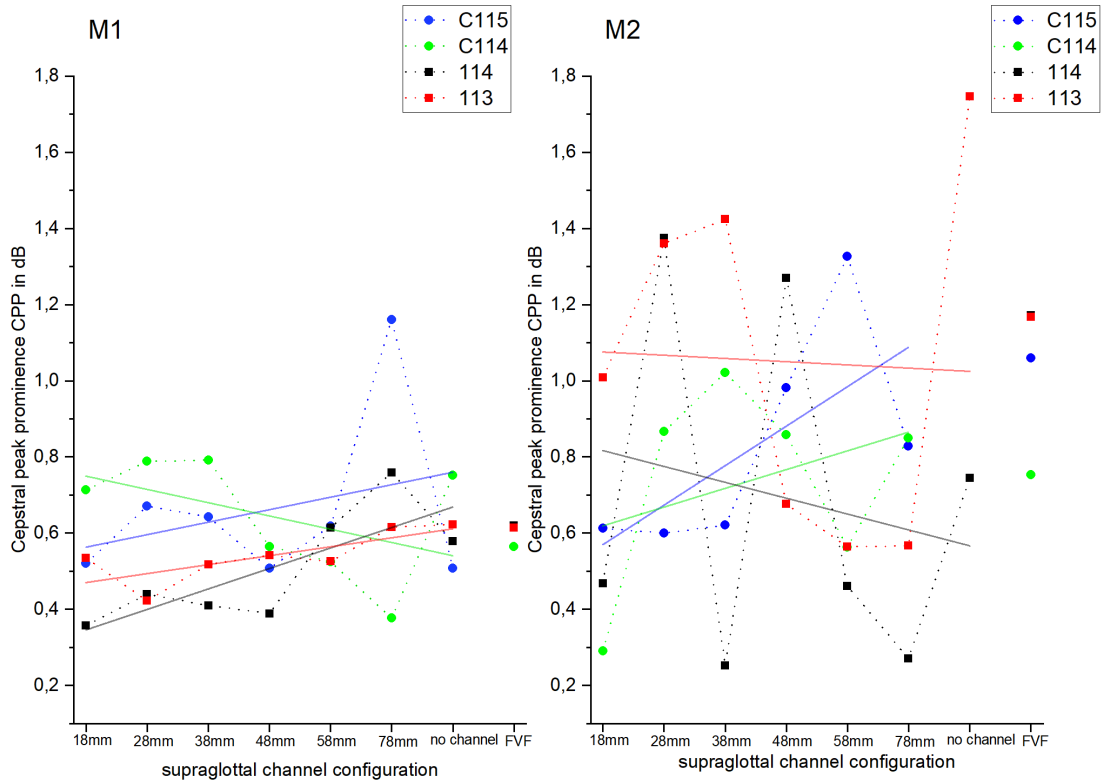


Figure 26 - Cepstral peak prominence [in dB] in M1 (left diagram) and M2 (right diagram) for different supraglottal channel configurations.

The values of the different model types are very similar in M1, whereas in M2 the trends of the CPP of the models differ among the models.

In M2 the general trend of the single-layer models differs to the multi-layer models: the values of the single-layer models decrease during the enlargement of the supraglottal channel, while the values of the multi-layer models increase.

The difference of CPP values within the multi-layer models in M1 in the *no channel* channel configuration and the *18mm* channel configuration is small.

In M1 and M2 the values of the single-layer models of the *no channel* channel configuration are higher than the values of the *18mm* channel configuration.

In M1 the values of the *FVF* channel configurations are very close[10]. The values of the *FVF* channel configuration in M2 are all higher than in the *18mm* channel configuration. This seems to indicate a higher tonality and periodicity in M2 with the FVFs than without.

[10] Exact data M1: C115 0,56 dB; 114 0,62 dB; 113 0,61 dB

The values of the physiological measurement [51] and the values of the synthetic vocal fold models are difficult to compare quantitatively because the CPP is normally evaluated based on acoustic data. The sound of a subject and the synthetic models couldn't be compared because of the partly missing vocal tract of the models. Additionally, the differences in the calculation and qualitative deviation of the CPP make a comparison of different sources complicated [58].

In the following, Figure 27 illustrates the results of the synthetic vocal folds of the cepstral peak prominence in dB and the realistic reproduction of physiological acoustics in comparison to the oscillations of human vocal folds. This is based on clinically relevant and objective parameters of the high-speed videos of the synthetic vocal fold movement.

Subsequently, the data by Monnappa and Balasubramaniam and the data of the synthetic larynx models C115, C114, 114 and 113 will be compared according to each parameter [52] [5]. Additionally, the different gender (female and male [52]) are subdivided to gain more information about the behave of the synthetic vocal fold models.

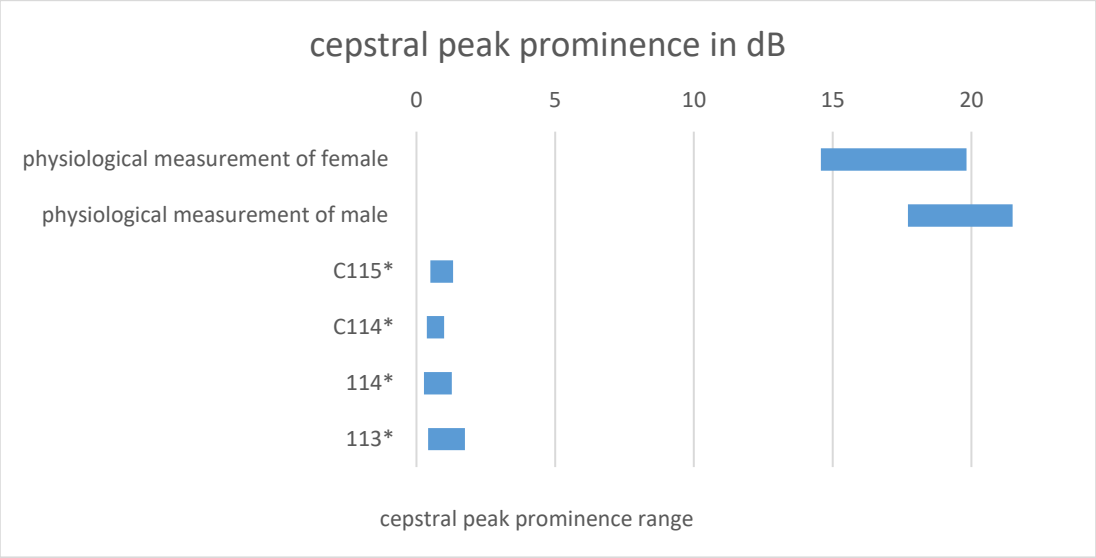


Figure 27 - Synthetic vocal fold model data in comparison to the physiological human data: Ranges of the cepstral peak prominence in dB (Physiological female and male versus 113, 114, C114 and C115 model) [52] [5].

As displayed in Figure 27 the range of the CPP of the physiological measurements differ to the ones of the vocal fold models. This might be caused by the origin of the CPP, which is within the acoustics and is therefore not comparable to the vocal fold models based on deviations in their composition.

3.8 Maximum area declination rate

The next two tables (Table 19 and Table 20) show the mean and the range values of all supraglottal channel configurations without FVF for the maximum area declination rate. The tables are subdivided in the M1 and M2, as well as the different vocal fold models.

parameter	**mean (all configurations without FVF)**							
	M1				M2			
	113	114	C114	C115	113	114	C114	C115
maximum area declination rate	-268	-325,1	-163,9	-129,1	-802,5	-643,8	-782	-465,9

Table 19 - Mean – Maximum area declination rate of all channel configurations without FVF in M1 and M2.

In Table 19 the mean shows significantly higher negative values within M2 than M1. The mean values of the single-layer models are higher negative than the ones of the multi-layer models. It is to say, that the single-layer models in M2 show a greater harmonic, which may be caused by the incomplete glottal closure of the multi-layer models during M2 in comparison to the single-layer models.

parameter	**range (all configurations without FVF)**							
	M1				M2			
	113	114	C114	C115	113	114	C114	C115
maximum area declination rate	186,2	119,6	254,5	230,5	311,8	171,9	798,1	209

Table 20 - Range (difference between the maximum value and the minimum value) – maximum area declination rate of all channel configurations without FVF in M1 and M2.

The range, which is displayed in Table 20, show in general higher values in M2 than in M1, except for C115. In M1 the single-layer models have a smaller range than the multi-layer models.

In the following, Figure 28 illustrates the maximum area declination rate for a better visual analysis of the vocal fold model movements. The different models C115, C114, 114 and 113 can be distinguished by different colors and signs of the single graphs.

In Figure 28, the x-axis shows the supraglottal channel configurations from 18mm up to 78mm, no channel and the supraglottal channel with FVFs. The y-axis shows the maximum area declination rate.

All models and channel configurations are recorded in two different oscillation modes. M1 (left side of the diagram) is without, while M2 (right side of the diagram) is with contact between the vocal folds in each oscillation cycle.

The data points represent the mean over all cycles of the detected high-speed recordings, while the standard deviation is displayed in the whiskers of the boxplot-like display.

The solid lines show the trend line of the respective model, while the dotted lines represent linear devolution between the single channel configurations. The *FVF* channel configuration is excluded from the calculation of the regression line.

In M2, data is missing for the multi-layer models in the *FVF* channel configuration. This is because M2 did not occur in this channel configuration with multi-layer models. The data for M1 with model C114 and with FVF are missing as well [5, p. 61].

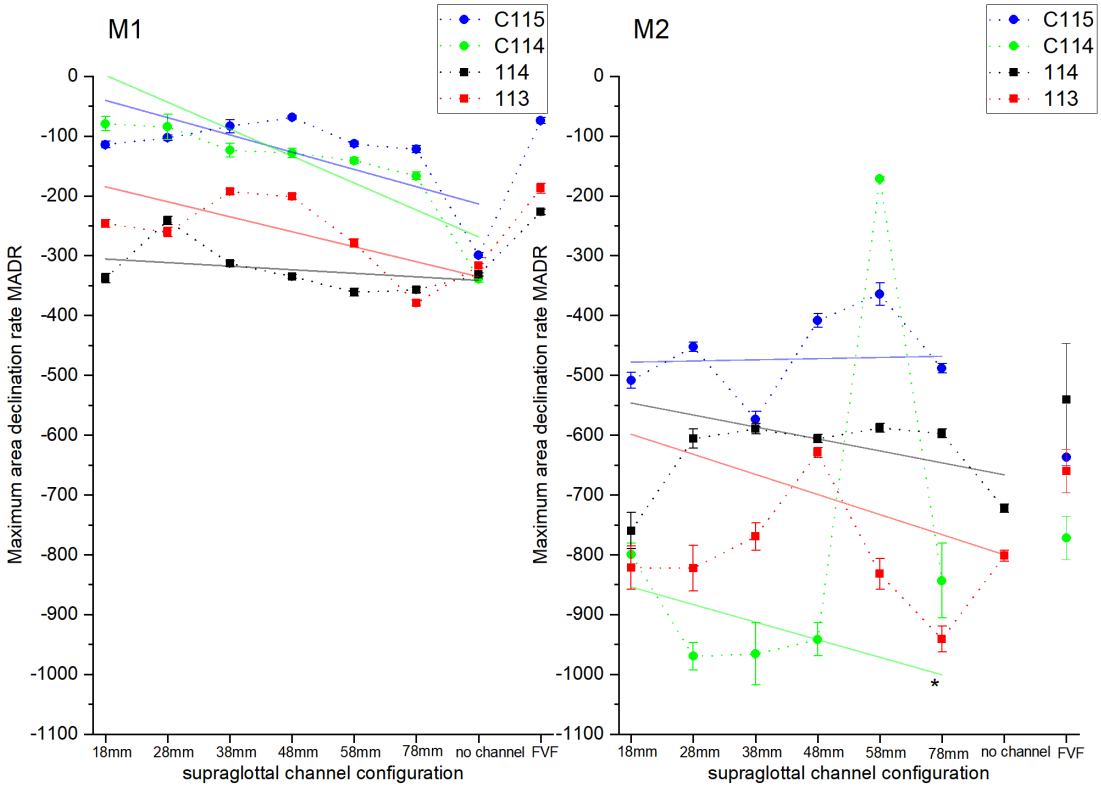


Figure 28 –Mean maximum area declination rate and standard deviation in M1 (left diagram) and M2 (right diagram) for different supraglottal channel configurations.

The general trend during the change of the diameter of the supraglottal channel is decreasing in both oscillation modes and for all models except for C115 in M2, which shows a constant trend. In general, the enlargement of the supraglottal channel leads to a faster closure of the glottal area.

The values in M1 are higher than the ones in M2, but in M2 the range of values is wider within the all models. Also, the standard deviation of all model types is greater in M2 than in M1. The higher MADR in M1 indicates a slower decrease of the glottal area during the closing motion, than in M2.

113 and 114 have lower values in M1 than C114 and C115. This seems to indicate a faster closing motion of the single-layer models.

In M1 113 shows higher values than 114, which changes in M2 where the values of 114 are higher than those of 113.

C115 has higher values than C114 in M1 and M2, which shows a slower closing of the glottal area within C115.

In M1 the values of the *no channel* channel configuration are close to each other[11], while the values of the *18mm* channel configuration are widely spread for the different models. In M1 and M2 the single-layer models show little difference for the *no channel* configuration and the *18mm* channel configuration.

In M1 all values of the *FVF* channel configuration are higher than the values of the *18mm* channel configuration, the same correlation is displayed within the single-layer models in M2. This correlation shows that with the FVFs the closure of the glottal area is slower than without.

In the following, Figure 29 shows the results of the synthetic vocal folds of the maximum area declination rate and the realistic reproduction of physiological acoustics in comparison to the oscillations of human vocal folds. This is based on clinically relevant and objective parameters of the high-speed videos of the synthetic vocal fold movement.

Subsequently, the data by Schützenberger et al., Patel et al. and the data of the synthetic larynx models C115, C114, 114 and 113 will be compared regarding to each parameter [48] [53] [5]. Additionally, the different gender (female [48] and male [53]) are subdivided to gain more information about the behave of the synthetic vocal fold models.

[11] Exact data
M1: C115 -299,176 C114 -339,022
114 -331,5 113 -316,933

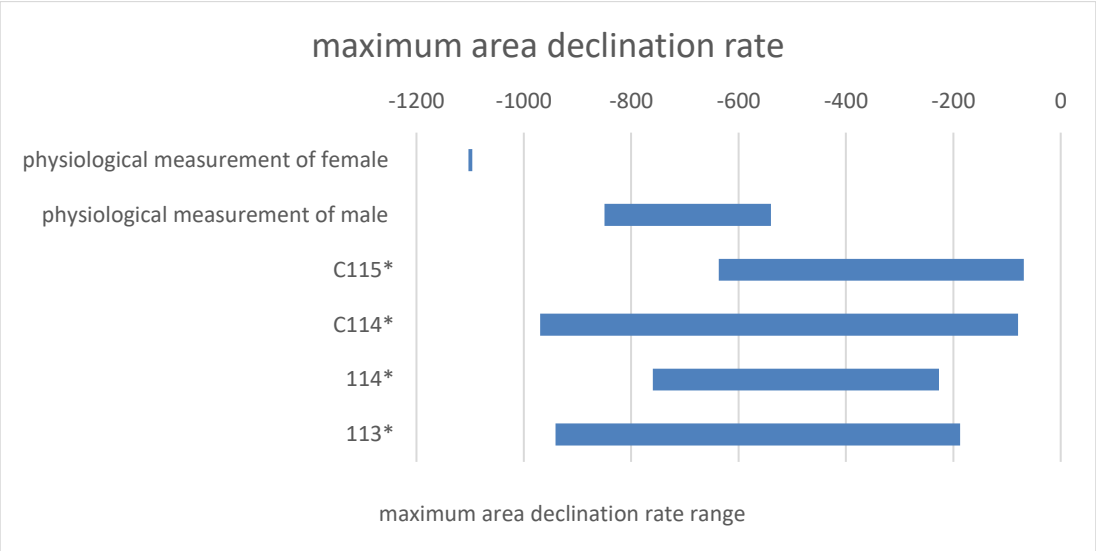


Figure 29 - Synthetic vocal fold model data in comparison to the physiological human data: Ranges of the maximum area declination rate (Physiological female and male versus 113, 114, C114 and C115 model). [48] [53] [5]

The exact value of the physiological measurement displays the same magnitude as the values of the single- and multi-layer models in Figure 29. The range of the models is much wider than the physiological [48] [53] ranges and show a higher numeric value.

The measured values of the male subjects [53] do overlap with those of the synthetic vocal folds. The synthetic vocal folds correlate with male subjects better than with female subjects.

4 Conclusion

The measurements of the single-layer and multi-layer vocal fold models which also differ in two types of stiffness (the models 113 and C114 are less stiff than the models 114 and C115) and in two different oscillation modes (M1 and M2)) are recorded by a high-speed camera located at the end of the supraglottal channel. The supraglottal channel itself features an extendable lateral diameter from 18mm to 78mm or no channel, as well as an insertable false vocal fold (FVF). The records during M1 show no contact between the vocal folds and thereby no total closure of the glottal gap during the oscillation, while in M2 the vocal folds close the glottal gap completely.

During the recordings the subglottal pressure and the flow rate were adapted for each configuration to induce vocal fold oscillations (see chapter 2.3 Measurement strategy). The analysis of the in total 64 high-speed videos and the resulting glottal area waveforms (GAW) is performed by GAT [46]. The values of seven objective parameter are extracted via Excel® and visualized with Orgin®.

The parameters are chosen by their significance in context of the process of phonation within humans (see chapter 2.4.2 Parameter choice with description): fundamental frequency, shimmer, mean-jitter, jitter, harmonics-to-noise ratio, cepstral peak prominence and maximum area declination. Symmetry parameters were not determined because of the exclusively symmetric oscillations of the silicone vocal folds.

To compare the GAW parameter of the multi-layer and single-layer models with the GAW parameters of physiological human phonation need no adjustment or conversion of the values. The data of the physiological GAW parameters originates from different papers and from both female or male subjects [49] [50] [48] [51] [53].

This leads to the central points of this dissertation: Do the values of the different parameters of the single-layer and multi-layer vocal fold models resemble the values of the physiological human phonation? Can they imitate the physiological oscillation of human vocal folds for research purposes? How are the synthetic models affected by the enlargement of the supraglottal channel, the stiffness of the silicon rubber, the contact of the vocal folds during the closing of the glottis or the insertion on FVFs?

The values of fundamental frequency of the male subjects [50], as well as the shimmer in percent [48], the harmonic to noise ratio (HNR) [48] and the maximum area declination rate (MADR) [53] of the male subjects match the values of the single- and multi-layer models. This overview demonstrates the similarity of the single- and multi-layer models with male subjects.

The values of the CPP show a large discrepancy because of the missing vocal tract above the supraglottal channel, which has an impact as well.

For a more detailed overview it is necessary to compare the individual vocal fold models and their configurations in this thesis referring to the GAW parameters:

During the enlargement of the supraglottal channel diameter there is a decrease of the values in the fundamental frequency, the shimmer in percent, the MADR and in the two jitter parameters. It can be assumed that the enlargement of the supraglottal channel diameter stabilizes the oscillation of the vocal folds. The change of the size of the supraglottal channel diameter has a higher effect on the oscillation of the multi-layer models than on the single-layer models, which can be seen in the greater range in of the multi-layer models.

The range of the stiffer models (114 and C115) is mostly lower than the one of the less stiff models, which indicates a more stable oscillation along with the stiffer models.

The two oscillation modes show different effects on the amplitude and the fundamental frequency. In M2 the range of the shimmer and fundamental frequency is smaller than in M1, which indicates a more stabilizing effect of M2 than M1 in all model types. The same correlation is demonstrated within the behavior of the jitter values. Referring to the noise quality the CPP shows a greater mean in M2 than in M1 in all model types, especially in the single-layer models. The single-layer model in M2 has higher CPP values than the multi-layer models, which can be attributed to the incomplete glottal closure of the multi-layer models in comparison to the complete closure of the single-layer models and its effect on the harmonics of the phonation.

The effect of the FVF can only be compared to the *18mm* channel configuration in all models, because the FVF are only mounted inside the *18mm* supraglottal channel. The values of the shimmer in percent are higher during the FVF are mounted in the supraglottal channel. The values of the CPP in M2 and the MADR show the same behavior. Inserting the FVF, the tonality and the periodicity are increased.

The use of the objective parameters as a diagnostic tool for patients with dysphonia is only possible with a given range of physiological values. This can be achieved by the intensive analysis of the parameters measured in realistic synthetic models or subjects to account for the diversity of human vocal folds.

To simplify the reproducibility and to determine a clear setting, the usage of synthetic vocal fold models is needed. Additionally, there should be a separation of female and male values. In this case, another model type is needed.

5 Bibliography

[1] H. Kutta, S. Knipping, H. Claassen and F. Paulsen, "Update Larynx: funktionelle Anatomie unter klinischen Gesichtspunkten," *HNO*, p. 583–598, 55 Apr 2007.

[2] H. Gray and W. Lewis, Anatomy of the Human Body, Philadelphia: Lea & Febiger, 1918.

[3] M. Schünke, E. Schulte, U. Schumancher, M. Voll and K. Wesker, Prometheus Lernatls der Anatomie - Kopf, Hals und Neuroanatomie, Band 3 Stuttgart: Thieme Verlag, 2012, p. 204.

[4] I. Titze, The Myoelastic-Aerodynamic Theory of Phonation, Denver: National Center for Voice and Speech, 2006.

[5] S. Kniesburges, Fluid-Structure-Acoustic Interaction during Phonation in a Synthetic Larynx Model, Aachen: Shaker Verlag, 2014.

[6] S. Silbernagl and R. Klinke, Lehrbuch der Physiologie, Stuttgart: Thieme, 2005.

[7] H. Pape, A. Kurtz and S. Silbernagl, Physiologie, Stuttgart: Thieme Verlag, 2014.

[8] S. Hammer, Stimmtherapie mit Erwachsenen, Heidelberg: Springer Medizin Verlag Heidelberg, 2005.

[9] J. Wendler, W. Seidner and U. Eysholdt, Lehrbuch der Phoniatrie und Pädaudiologie, Stuttgart: Thieme, 2005.

[10] M. Döllinger, J. Kobler, D. Berry, D. Mehta, G. Luegmair and C. Bohr, "Experiments on Analysing Voice Production: Excised (Human, Animal) and In Vivo (Animal) Approaches," *Current Bioinformatics*, pp. 286-304, Jun 2011.

[11] S. Prudil, "Dreidimensionale Stimmlippenbewegungen im exzidierten Schweinemodellkehlkopf," Phoniatrischen und Pädaudiologischen Abteilung der Friedrich-Alexander-Universität Erlangen-Nürnberg, Erlangen, 2014.

[12] S. Kniesburges, S. Thomson, A. Barney, M. Triep, P. Sidlof, J. Horácek, C. Brückner and S. Becker, "In Vitro Experimental Investigarion of Voice Production," *Current Bioinformatics*, pp. 305-322, 6 Sep 2011.

[13] Maytag, A, Robitaille, M, Rieves, A, J. Madsen, B. Smith and J. Jiang, "Use of the rabbit larynx in an excised larynx setup," *Journal of Voice,* pp. 24-28, 27 Jan 2013.

[14] F. Alipour and E. Finnegan, "On the acoustic effects of the supraglottic structures in excised larynges," *J Acoust Soc Am,* pp. 2984-2992, 133 (5) May 2013.

[15] P. Murray and S. Thomson, "Synthetic, Multi-Layer, Self-Oscillating Vocal Fold Model Fabrication," *Journal of Visualized Experiments,* 58 Dec 2011.

[16] B. Kucinschi, R. Scherer, K. DeWitt and T. Ng, "An experimental analysis of the pressures and flows within a driven mechanical model of phonation," *The Journal of the Acoustical Society of America,* p. 3011–3021, 119 May 2006.

[17] Q. Xue, X. Zheng, R. Mittal and S. Bielamowicz, "Subject-specific computational modeling of human phonation," *The Journal of the Acoustical Society of America,* pp. 1445-1456, 135 Mar 2014.

[18] X. Zheng, R. Mittal and S. Bielamowicz, "A computational study of asymmetric glottal jet deflection during phonation," *The Journal of the Acoustical Society of America* , p. 2133–2143, 129 Apr 2011.

[19] M. Döllinger, "The Next Step in Voice Assessment: High-Speed Digital Endoscopy and Objective Evaluation," *Current Bioinformatics,* pp. 101-111, Apr 2009.

[20] R. Wegel, "Theory of Vibration of the Larynx," *Nokia Bell Labs,* pp. 207-227, 9 Jan 1930.

[21] R. Scherer, K. DeWitt and B. Kucinschi, "The effect of exit radii on intraglottal pressure distributions in the convergent glottis," *The Journal of the Acoustical Society of America,* p. 2267–2269, 110 Oct 2001.

[22] M. Deverge, X. Pelorson, C. Vilain, P. Lagree, F. Chentouf, J. Willems and A. Hirschberg, "Influence of collision on the flow through in-vitro rigid models ofthe vocal folds.," *The Journal of the Acoustical Society of America,* p. 3354–3362, 114 Jan 2004.

[23] M. Triep, C. Brücker and W. Schröder, "High-speed PIV measurements of the flow downstream of a dynamic mechanical model of the human vocal folds," *Experiments in Fluids,* p. 232–245, 39 Jul 2005.

[24] S. Thomson, L. Mongeau and F. Frankel, "Physical and numerical flow-excited vocal fold model," *MAVEBA*, pp. 147-150, 2003.

[25] M. Spencer, T. Siegmund and L. Mongeau, "Determination of superior surface strains and stresses, and vocal fold contact pressure in a synthetic larynx model using digital image correlation," *The Journal of the Acoustical Society of America*, pp. 1089-1103, 123 Feb 2008.

[26] P. Murray, S. Thomson and M. Smith, "A synthetic, self-oscillating vocal fold model platform for studying augmentation injection," *Journal of Voice*, p. 133–143, 28 Mar 2014.

[27] J. Drechsel and S. Thomson, "Influence of supraglottal structures on the glottal jet exiting a two-layer synthetic, self-oscillating vocal fold model.," *The Journal of the Acoustical Society of America*, p. 4434–4445, 123 Jun 2008.

[28] P. Murray and S. Thomson, "Vibratory responses of synthetic, self-oscillating vocal fold models," *The Journal of the Acoustical Society of America*, p. 3428–3438, 132 Nov 2012.

[29] B. Pickup and S. Thomson, "Flow-induced vibratory response of idealized versus magnetic resonance imaging-based synthetic vocal fold models," *The Journal of the Acoustical Society of America*, pp. EL124-EL129, 128 Sep 2010.

[30] P. Šidlof, O. Doaré, O. Cadot and A. Chaigne, "Measurement of flow seperation in a human vocal folds model," *Experiments in Fluids*, pp. 123-136, 51 Jul 2011.

[31] S. Owaki, H. Kataoka and T. Shimizu, "Relationship Between Transglottal Pressure and Fundamental Frequency of Phonation - Study Using a Rubber Model," *Journal of Voice*, pp. 127-132, Mar 2010.

[32] H. Schutte, J. Svec and F. Sram, "First results of clinical application of videokymography," *Laryngoscope*, pp. 1206-1210, 108 Aug 1998.

[33] M. Fukahori, S. Chitose, K. Sato, S. Sueyoshi, T. Kurita, H. Umeno, Y. Monden and R. Yamakawa, "Regeneration of vocal fold mucosa using tissue-engineered structures with oral mucosal cells," *PLoS ONE*, pp. 1-15, 11 e0146151 Jan 2016.

[34] C. Krausert, A. Olszewski, L. Taylor, J. McMurray, S. Dailey and J. Jiang, "Mucosal wave measurement and visualization techniques," *Journal of Voice*, pp. 395-405, 25 Jul 2011.

[35] J. Lohscheller, H. Toy, F. Rosanowski, U. Eysholdt and M. Döllinger, "Clinically evaluated procedure for the reconstruction of vocal fold vibrations from endoscopic digital high-speed videos," *Medical Image Analysis*, pp. 400-413, 11 Aug 2007.

[36] B. Farnsworth, "High speed motion pictures of the human vocal cords," *Bell Laboratory Records*, pp. 203-208, 1940.

[37] M. Hess, H. Herzel, O. Köster, F. Scheurich and M. Gross, "Endoskopische Darstellung von Stimmlippenschwingungen: Digitale Hochgeschwindigkeitsaufnahmen mit verschiedenen Systemen," *HNO*, p. 685–693, 44 Dec 1996.

[38] P. Moore and H. von Leden, "Dynamic Variations of the Vibratory Pattern in the Normal Larynx," *Folia phonitatrica et logopaedia*, p. 10(4):205–238, Oct 1958.

[39] R. Patel, D. Dubrovskiy and M. Doellinger, "Characterizing vibratory kinematics in children and adults with high-speed digital imaging," *Journal of Speech Language and Hearing Research*, pp. 674-686, 57 Apr 2014.

[40] M. Echternach, S. Dippold and B. Richter, "High speed imaging using rigid laryngoscopy for the analysis of register transitions in professional operatic tenors," *Logopedics Phoniatrics Vocology*, pp. 1-8, 2014.

[41] J. Lohscheller, M. Doellinger, M. Schuster, R. Schwarz, U. Exsholdt and U. Hoppe, "Quantitative investigation of the vibration pattern of the substitute voice generator," *IEEE Transactions on Biomedical Engineering*, pp. 1394-1400, 51 Aug 2004.

[42] S. Zacharias, C. Myer, J. Meinzen-Derr, L. Kelchner, D. Deliyski and A. de Alarcón, "Comparison of Videostroboscopy and High-speed Videoendoscopy in Evaluation of Supraglottic Phonation," *Annals of Otology, Rhinology & Laryngology*, pp. 829-837, 125 Oct 2016.

[43] S. Zacharias, D. Deliyski and T. Gerlach, "Utility of Laryngeal High-speed Videoendoscopy in Clinical Voice Assessment," *Journal of Voice*, pp. 216-220, 32 Mar 2018.

[44] J. Rubin, R. Sataloff and G. Korovin, Diagnosis and treatment of voice disorders, San Diego: Plural Publishing, 2014.

[45] S. Riedlmeier, "Auslegung, Konstruktion und Analyse eines Schalldämpfers in einem Kanal zu Untersuchung künstlicher Stimmlippen," Institut of Process Machinery and Systems Engineering, University Erlangen-Nürnberg, Erlangen, 2010.

[46] Department for Phoniatrics and Pedaudiology, *Glottis Analysis Tools User guide,* Bohlenplatz 21, 91054 Erlangen, Germany: Department for Phoniatrics and Pedaudiology, 2014.

[47] E. Holmberg, R. Hillman and J. Perkell, "Glottal airflow and transglottal air pressure measurements for male and female speakers in soft, normal, and loud voice," *The Journal of the Acoustical Society of America,* no. 85 (4), pp. 511-529, 84 Aug 1988.

[48] A. Schützenberger, M. Kunduk, M. Döllinger, C. Alexiou, D. Dubrovskiy, M. Semmler, A. Seger and C. Bohr, "Laryngeal High-Speed Videoendoscopy: Sensitivity of Objective Parameters towards Recording Frame Rate," *BioMed Research International,* pp. 1-19, Jul 2016.

[49] K. Ahmad, Y. Yan and D. Bless, "Vocal fold vibratory characteristics in normal female speakers from high-speed digital imaging.," *Journal of Voice,* pp. 26(2):239-253, Mar 2012.

[50] S. Warhurst, P. McCabe, R. Heard, E. Yiu, G. Wang and C. Madill, "Quantitative Measurement of Vocal Fold Vibration in Male Radio Performers and Healthy Controls Using High-Speed Videoendoscopy," *PLoS ONE,* p. 9(6):e101128, Sep 2014.

[51] Y. Heman-Ackah, R. Sataloff, G. Laureyns, D. Lurie, D. Michael, R. Heuer, A. Rubin, R. Eller, S. Chandran, M. Abaza, K. Lyons, V. Divi, J. Lott, J. Johnson and J. Hillenbrand, "Quantifying the cepstral peak prominence, a measure of dysphonia," *Jornal of Voice,* pp. 783-788, 28 Nov 2014.

[52] D. Monnappa and R. Balasubramaniam, "Cepstral analysis of voice in healthy aged individuals," *Jornal of Laryngology and Voice,* vol. 2, no. 34-38, pp. 34-38, May 2015.

[53] R. Patel, D. Dubrovskiy and M. Döllinger, "Measurement of Glottal Cycle Characteristics Between Children and Adults: Physiological Variations," *Journal of Voice,* pp. 476-486, 28 Dec 2013.

[54] T. Lin, J. Chen, C. Liu, C. Lee, Y. Tsou and C. Chuang, "A feasibility study on non-invasive oxidative metabolism detection and acoustic assessment of human vocal cords by using optical technique," *Scientfic reports*, 2017.

[55] M. Brockmann, M. Drinnan, C. Storck and P. Carding, "Reliable Jitter and Shimmer Measurements in Voice Clinics: The Relevance of Vowel, Gender, Vocal Intensity, and Fundamental Frequency Effects in a Typical Clinical Task," *Journal of Voice*, vol. No. 1, no. 44-53, pp. 44-53, 25 Jan 2011.

[56] A. Laukkanen, I. Ilomäki, K. Leppänen and E. Vilkman, "Acoustic Measures and Self-reports of Vocal Fatigue by Female Teachers," *Journal of Voice*, pp. 283-289, 22 May 2008.

[57] V. Birk, S. Kniesburges, M. Semmler, D. Berry, C. Bohr, M. Döllinger and A. Schützenberger, "Influence of glottal closure on the phonatory process in ex vivo porcine larynges," *The Journal of the Acoustical Society of America*, vol. 4, no. Oct, pp. 2197-2207, 142 Oct 2017.

[58] V. Birk, A. Sutor, M. Döllinger, C. Bohr and S. Kniesburges , "Acoustic Impact of Ventricular Folds on Phonation Studied in Ex Vivo Human Larynx Models," *Acta Acustica united with Acustica*, vol. Volume 102, no. Number 2, pp. 244-256, Mar 102 2016.

[59] J. Gliński, J. Horabik and J. Lipiec, Encyclopedia of Agrophysics, Springer, 2011.

[60] U. Harten, Physik für Mediziner, Heildelberg: Springer Verlag, 2014.

6 Table of figures

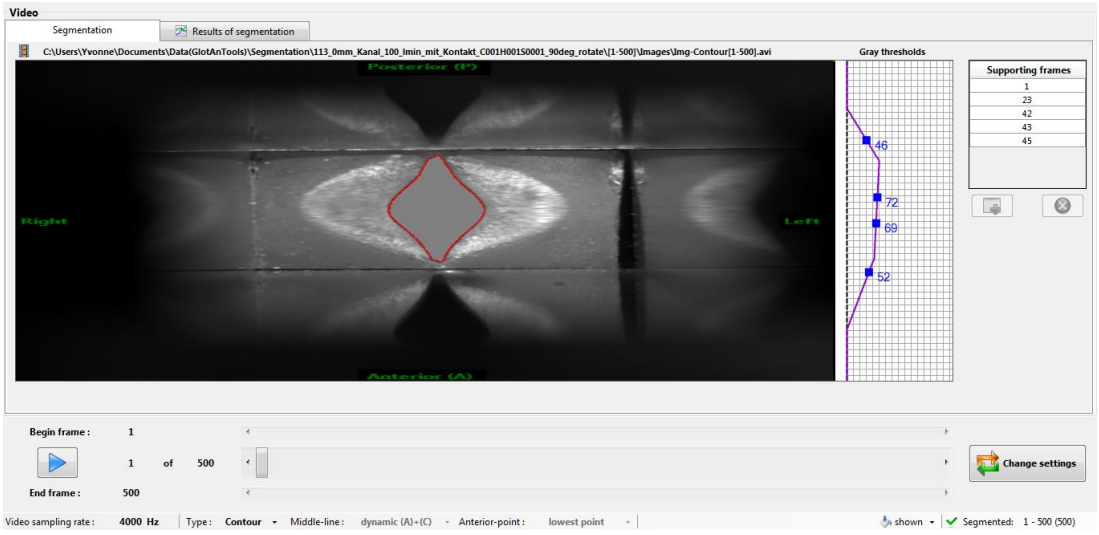
6.1 Figures

6.2 Tables

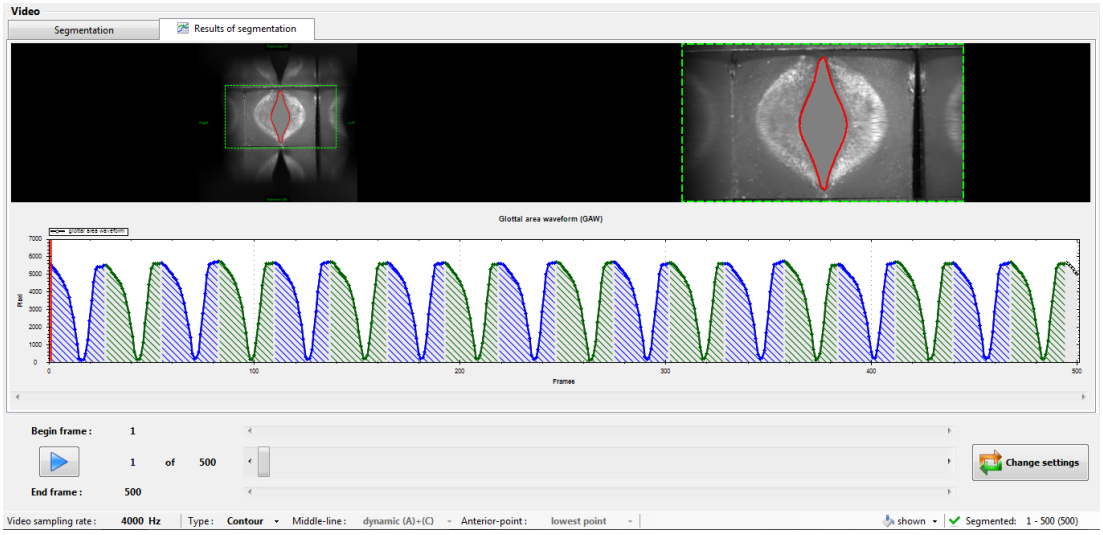
7 Nomenclature

CPP	cepstral peek prominence
D	diameter
Fps	frames per second
FVF	false vocal folds
GAT	Glottis analysis tool
GAW	Glottal area waveform
HSV	high-speed videoendoscopy
MADR	maximum area declination rate
MRI	magnetic resonance imaging
M1	oscillation mode 1
M2	oscillation mode 2

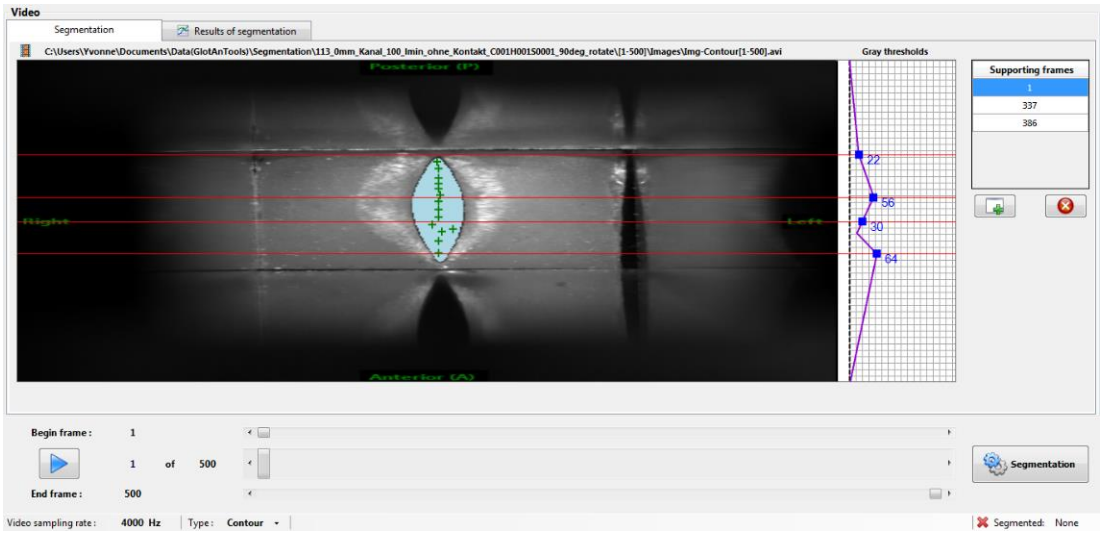
8 Appendix



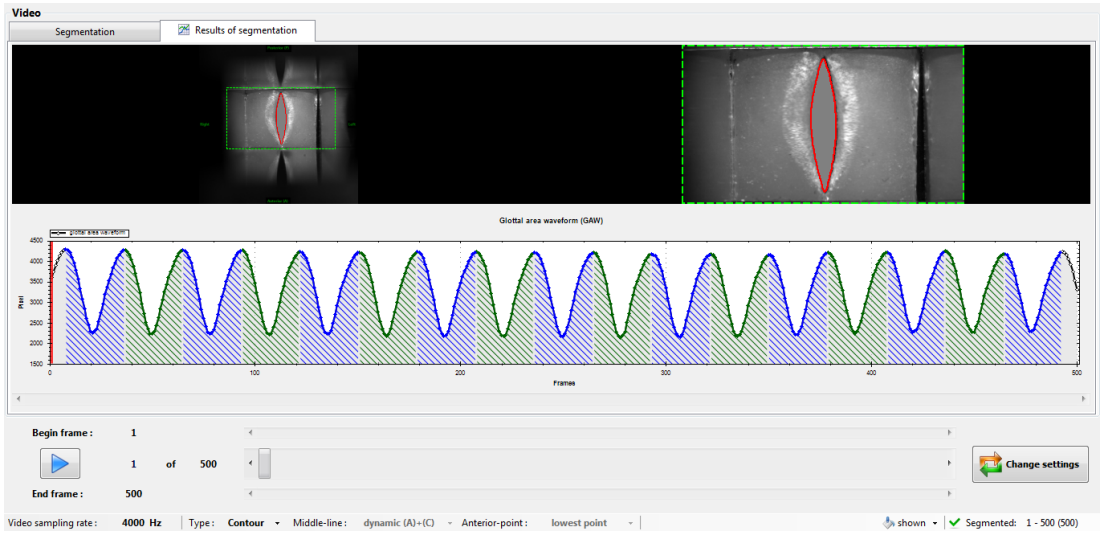
A Figure 1 - Model 113 in 18mm channel configuration in M2 during the setting of the maximum glottal gap with seatpoints with GAT.



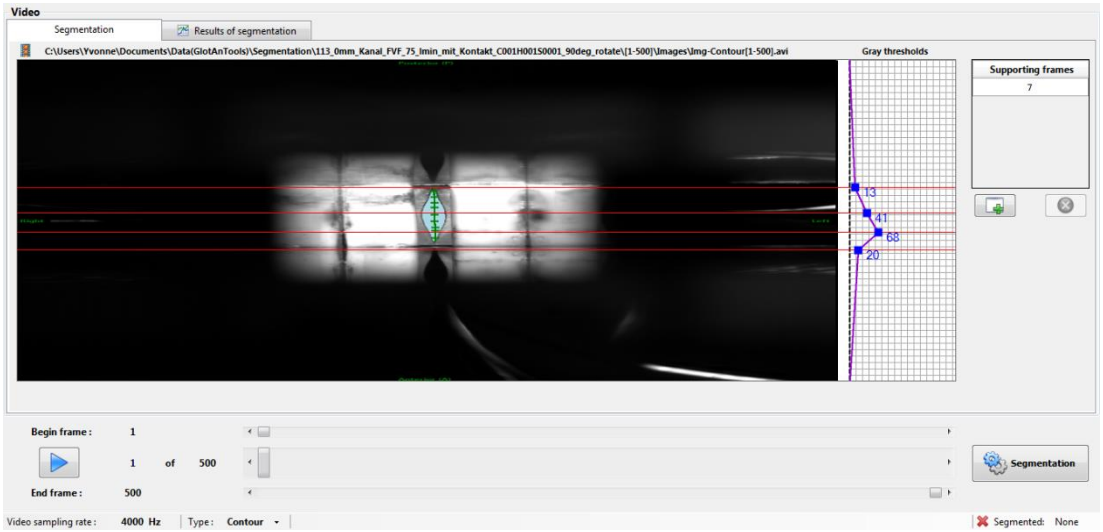
A Figure 2 - Segmentation of the model 113 in 18mm channel configuration M2 with GAT.



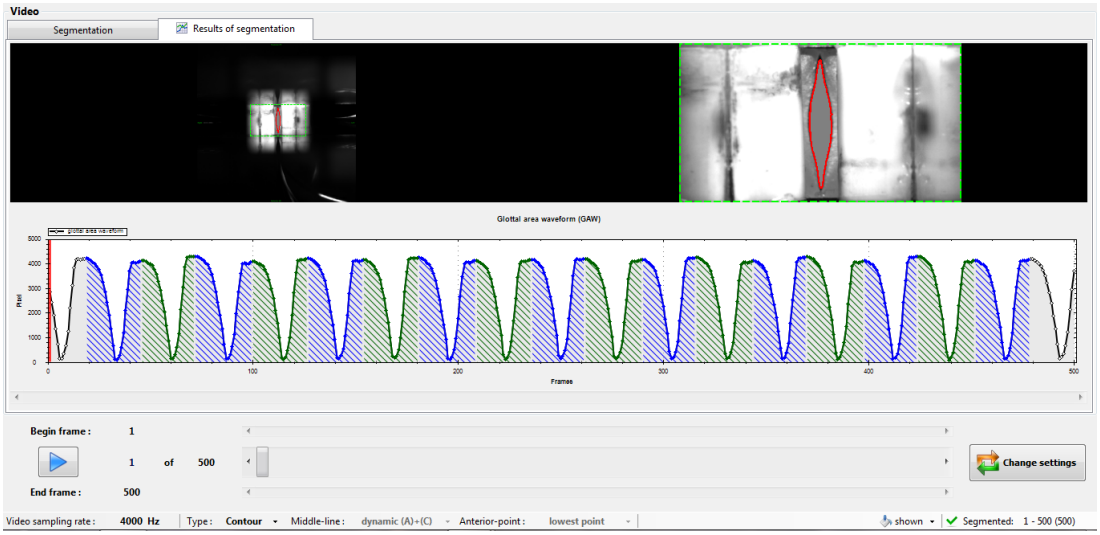
A Figure 3 - Model 113 in 18mm channel configuration in M1 during the setting of the maximum glottal gap with seatpoints with GAT.



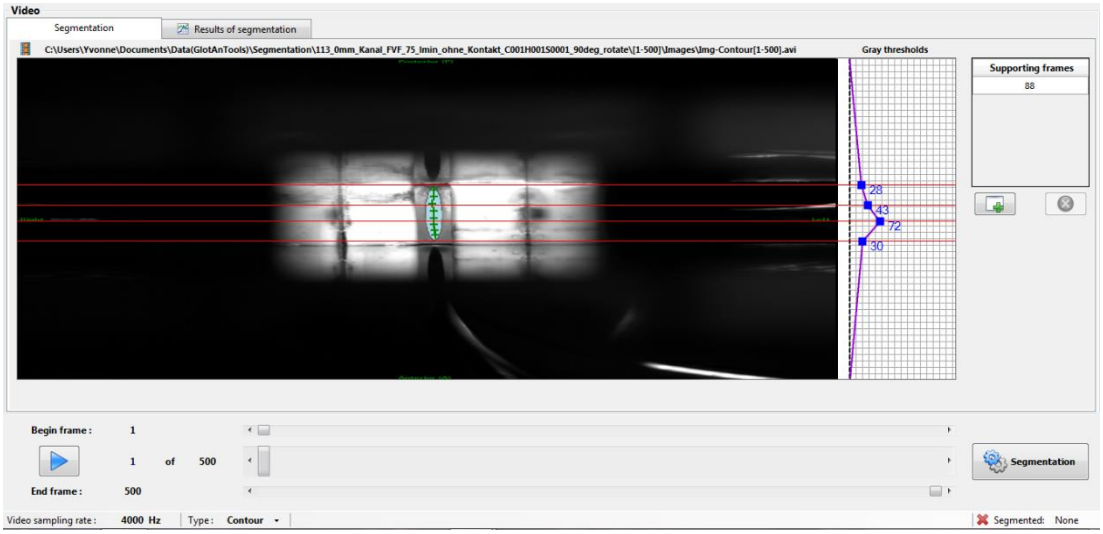
A Figure 4 - Segmentation of the model 113 in 18mm channel configuration M1 with GAT.



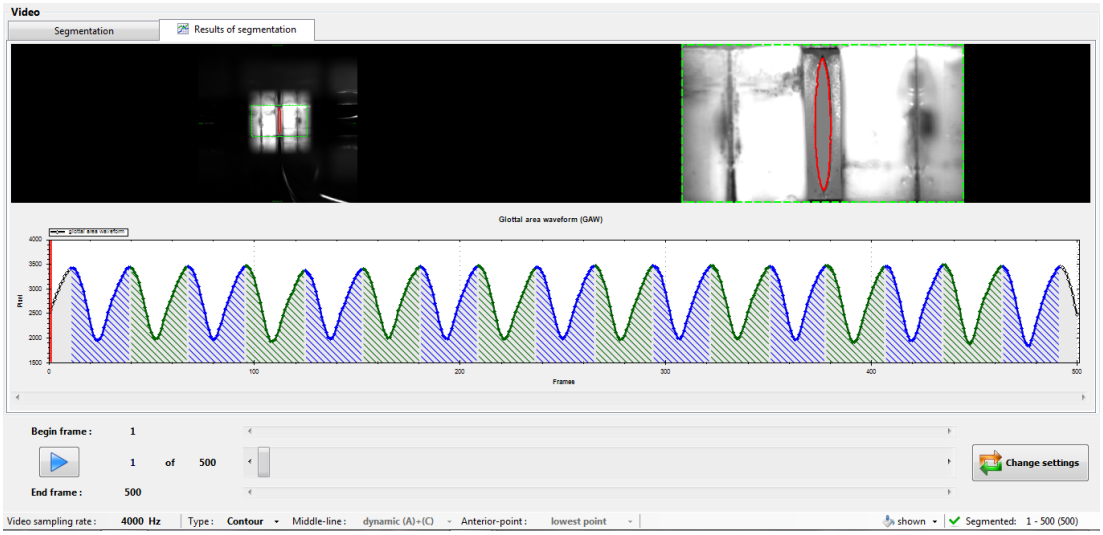
A Figure 5 - Model 113 in 18mm with FVF channel configuration in M2 during the setting of the maximum glottal gap with seatpoints with GAT.



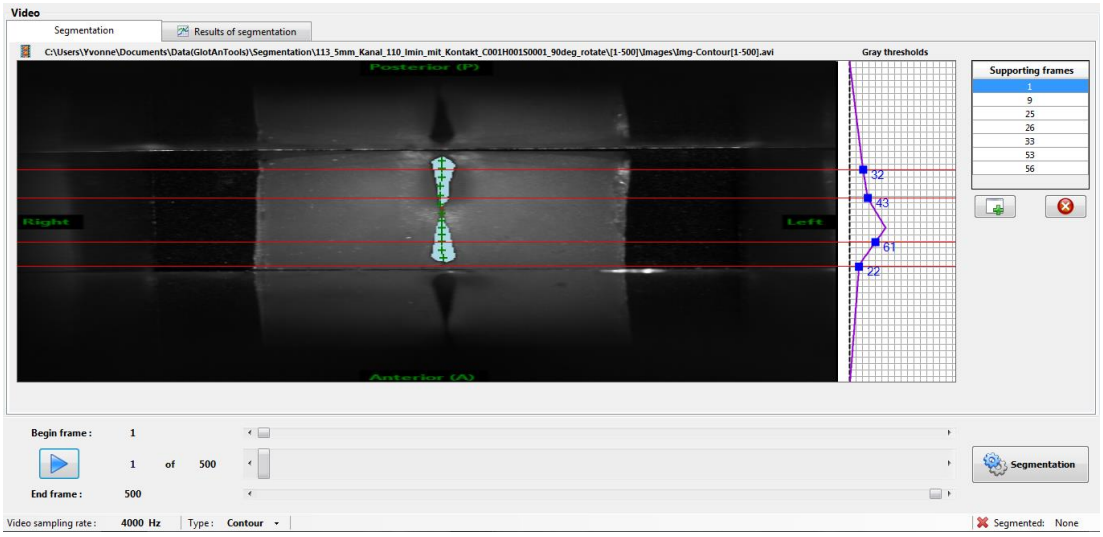
A Figure 6 - Segmentation of the model 113 in 18mm with FVF channel configuration M2 with GAT.



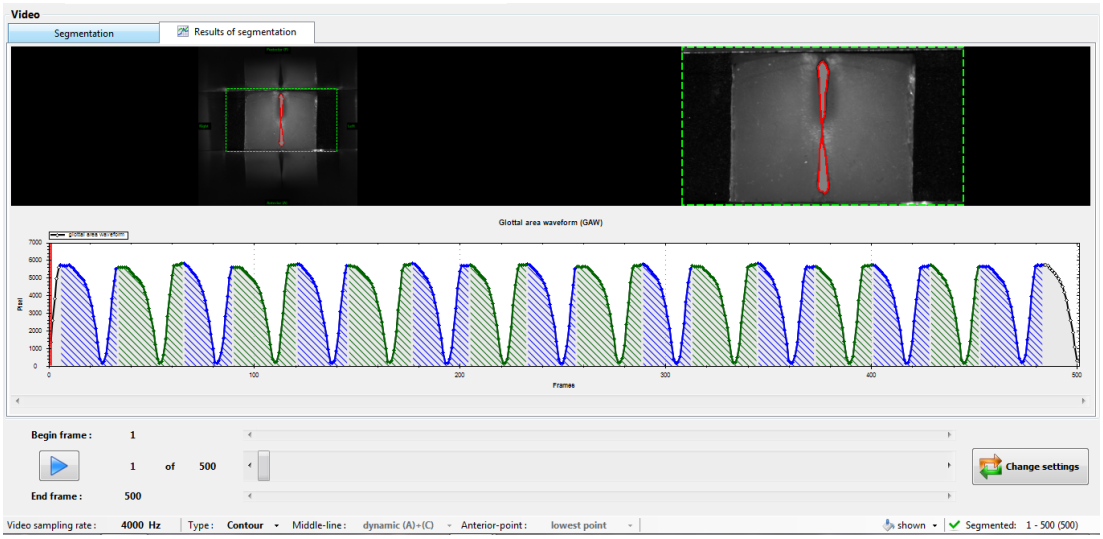
A Figure 7 - Model 113 in 18mm with FVF channel configuration in M1 during the setting of the maximum glottal gap with seatpoints with GAT.



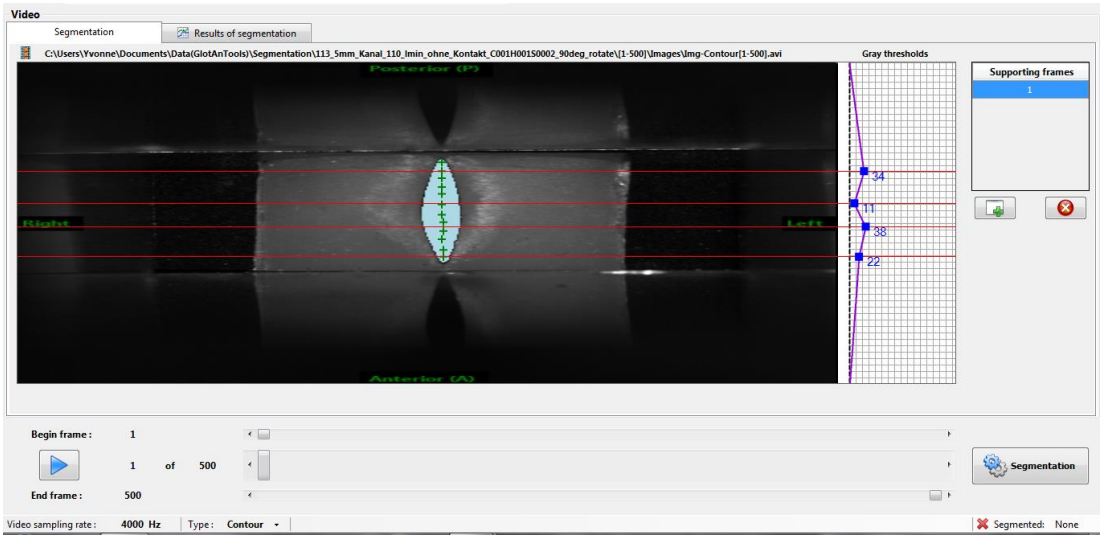
A Figure 8 - Segmentation of the model 113 in 18mm with FVF channel configuration M1 with GAT.



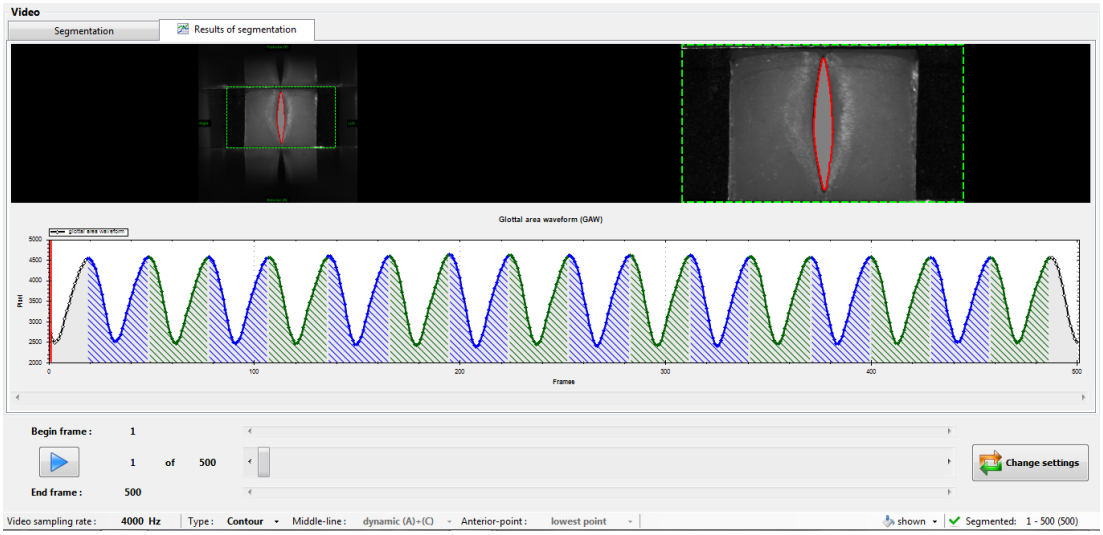
A Figure 9 - Model 113 in 28mm channel configuration in M2 during the setting of the maximum glottal gap with seatpoints with GAT.



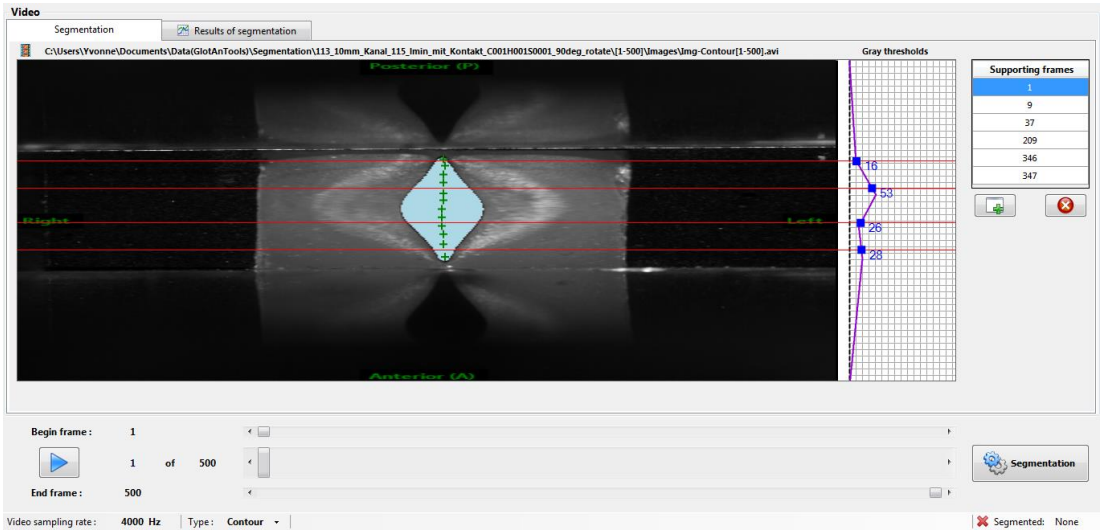
A Figure 10 - Segmentation of the model 113 in 28mm channel configuration M2 with GAT.



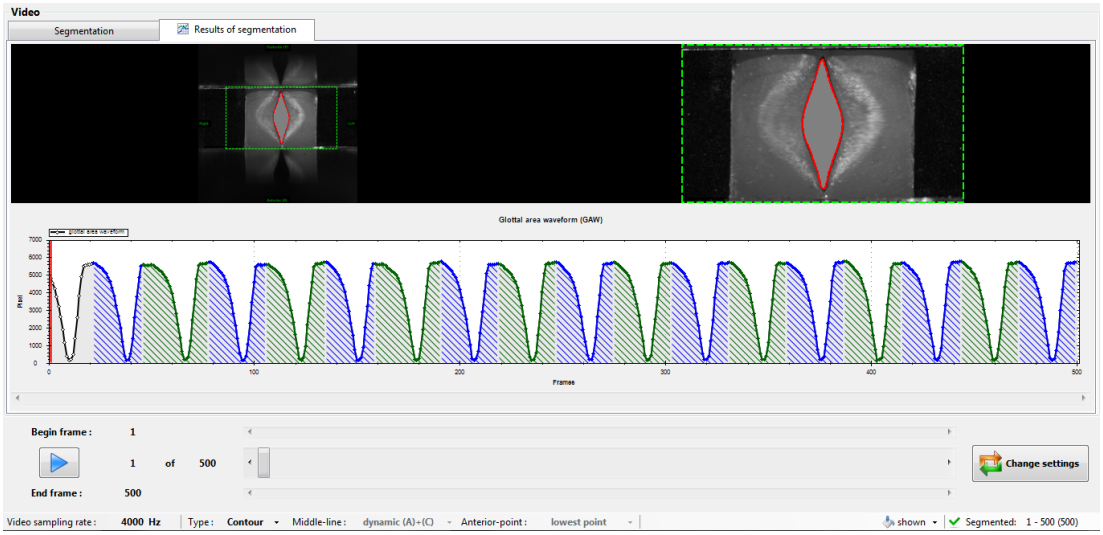
A Figure 11 - Model 113 in 28mm channel configuration in M1 during the setting of the maximum glottal gap with seatpoints with GAT.



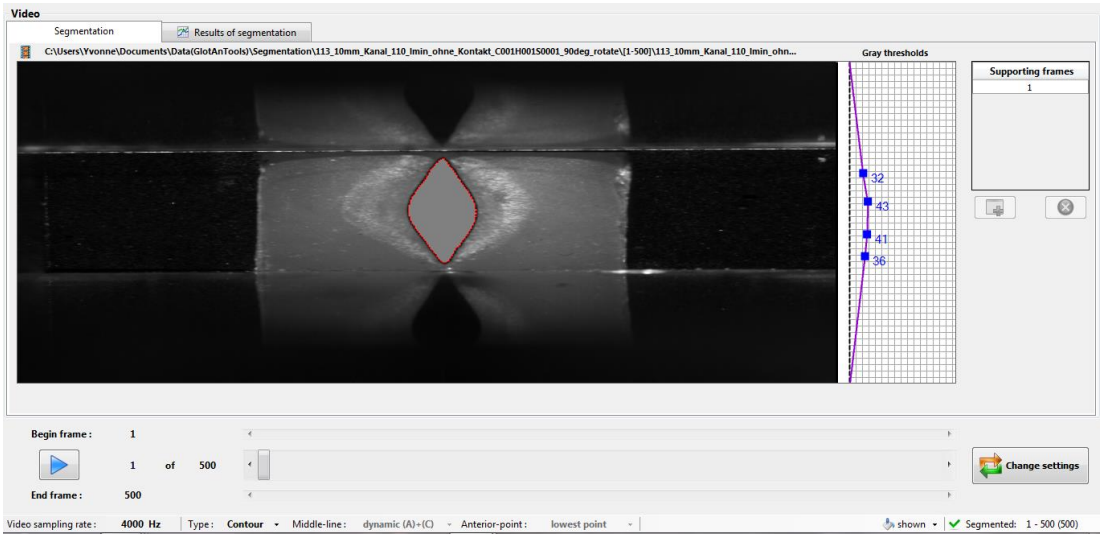
A Figure 12 - Segmentation of the model 113 in 28mm channel configuration M1 with GAT.



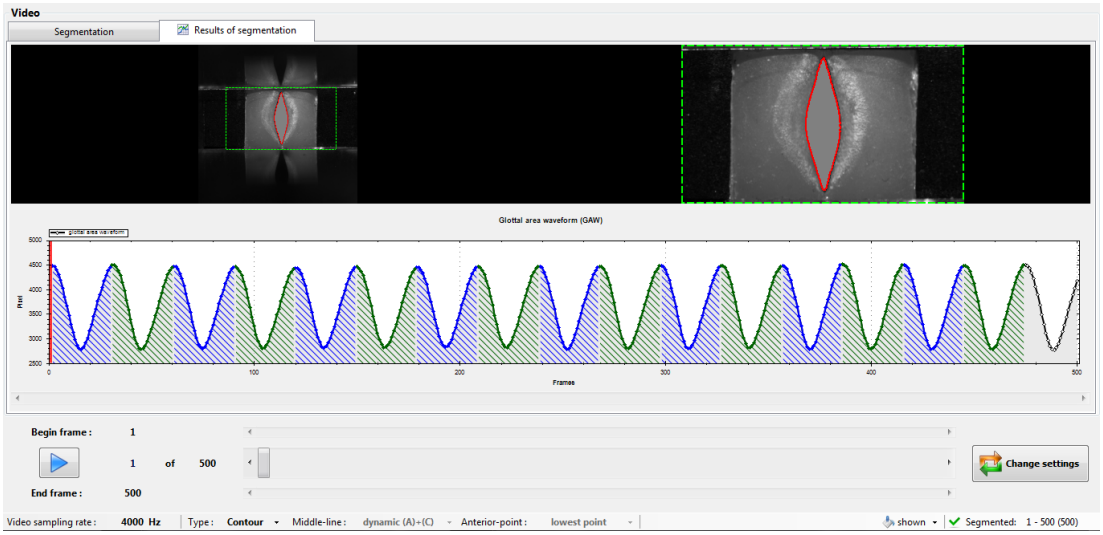
A Figure 13 - Model 113 in 38mm channel configuration in M2 during the setting of the maximum glottal gap with seatpoints with GAT.



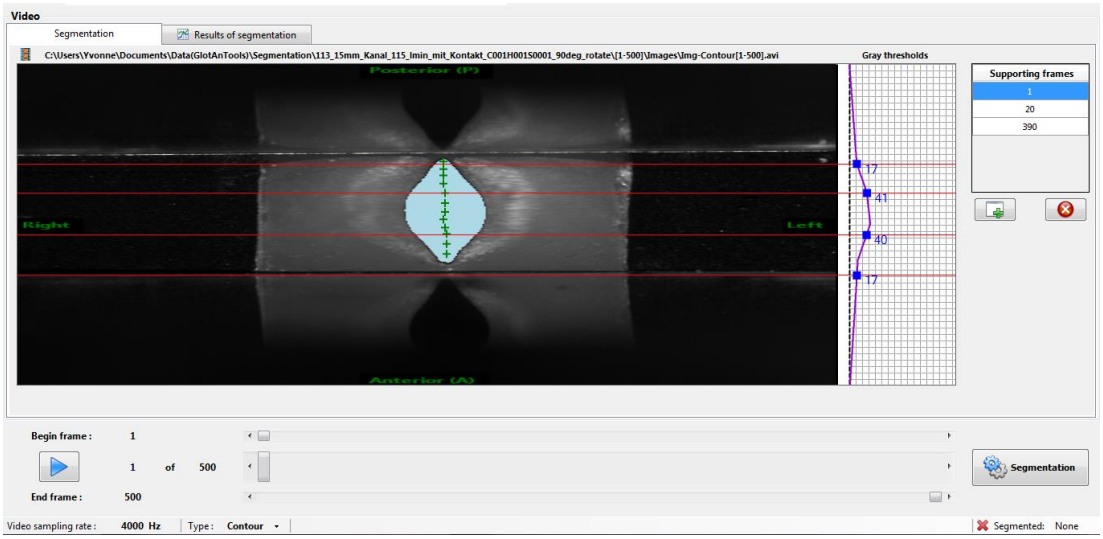
A Figure 14 - Segmentation of the model 113 in 38mm channel configuration M2 with GAT.



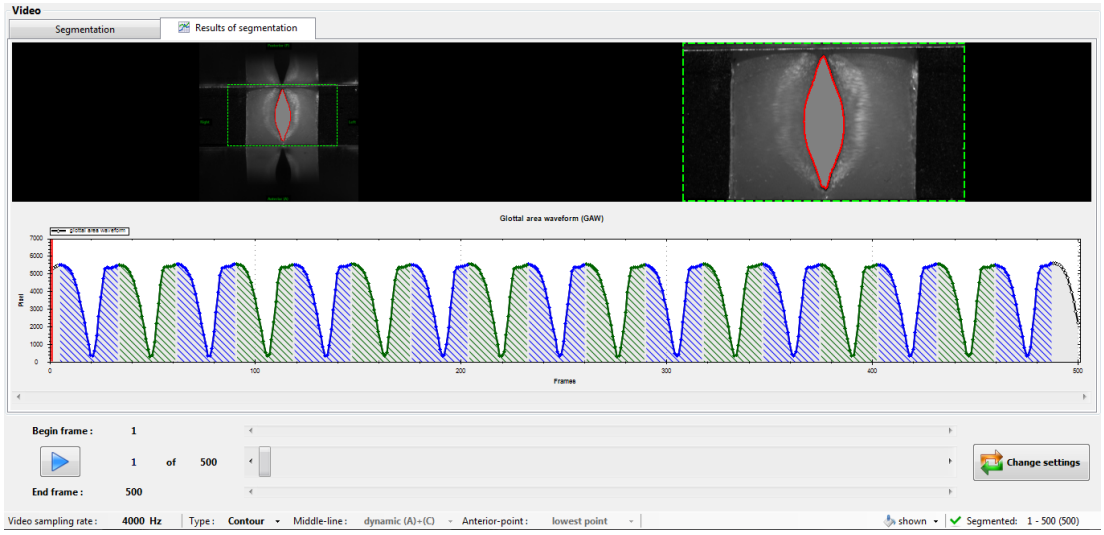
A Figure 15 - Model 113 in 38mm channel configuration in M1 during the setting of the maximum glottal gap with seatpoints with GAT.



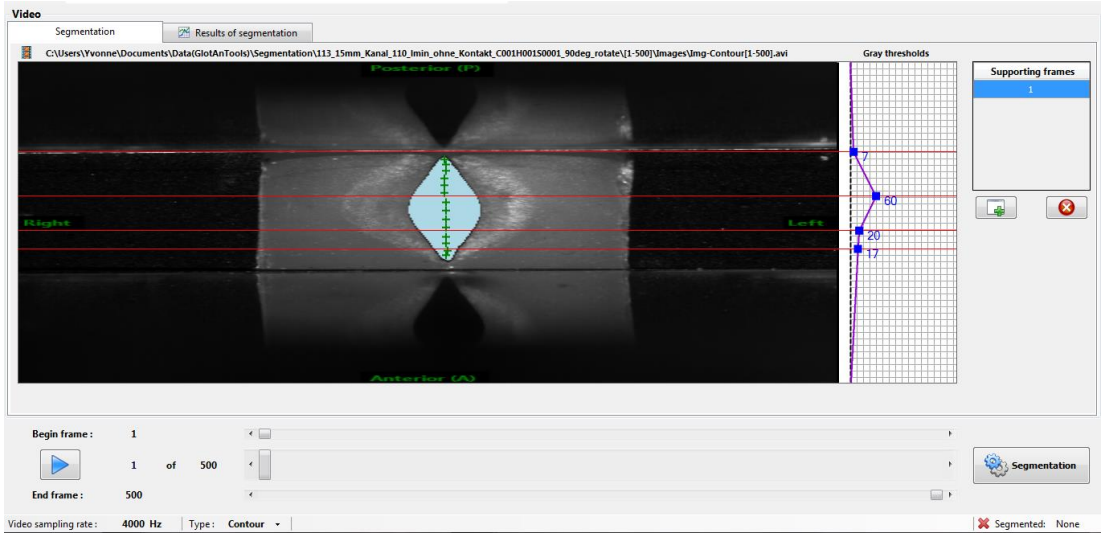
A Figure 16 - Segmentation of the model 113 in 38mm channel configuration M1 with GAT.



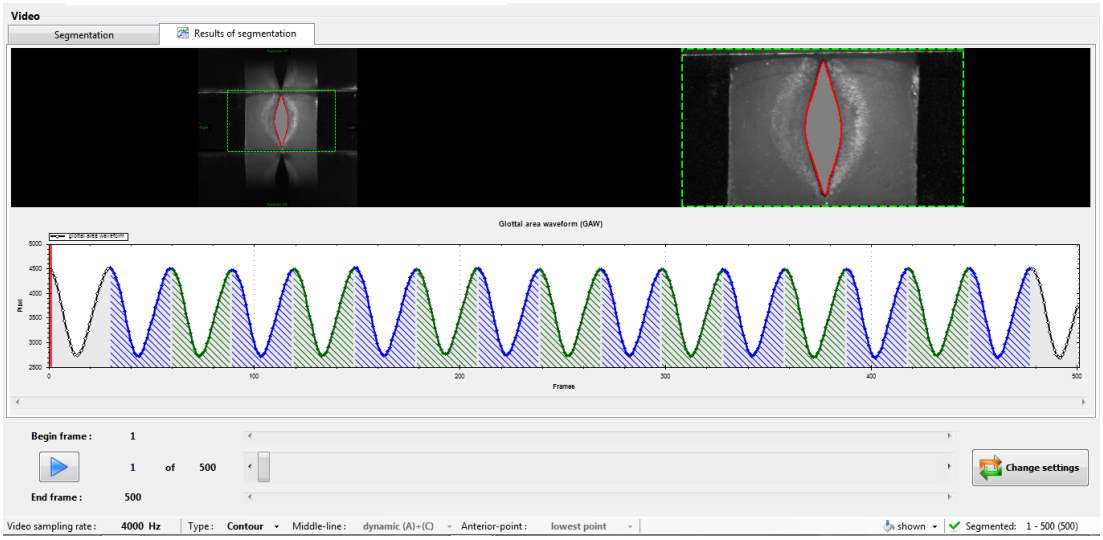
A Figure 17 - Model 113 in 48mm channel configuration in M2 during the setting of the maximum glottal gap with seatpoints with GAT.



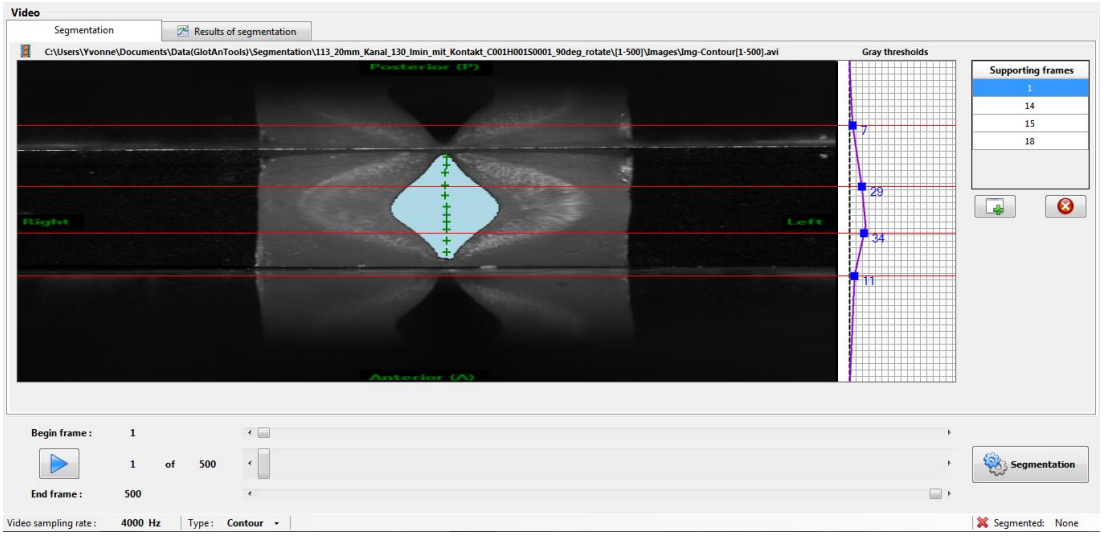
A Figure 18 - Segmentation of the model 113 in 48mm channel configuration M2 with GAT.



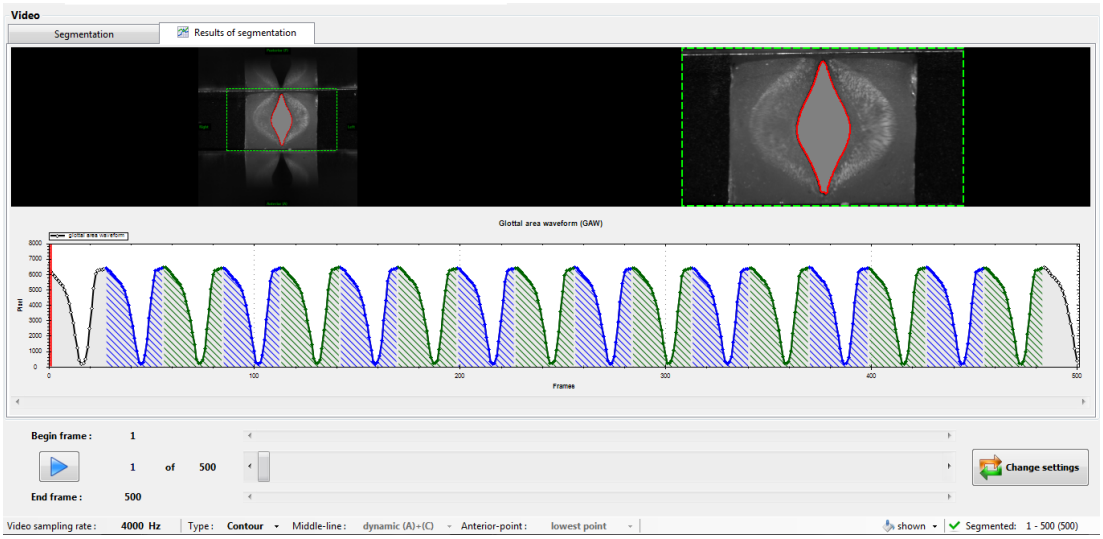
A Figure 19 - Model 113 in 48mm channel configuration in M1 during the setting of the maximum glottal gap with seatpoints with GAT.



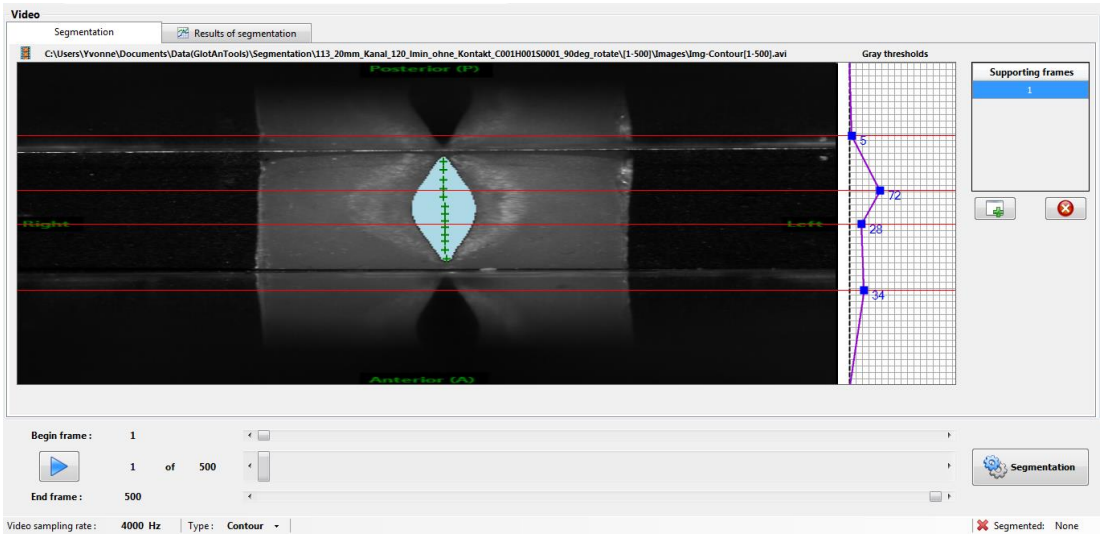
A Figure 20 - Segmentation of the model 113 in 48mm channel configuration M1 with GAT.



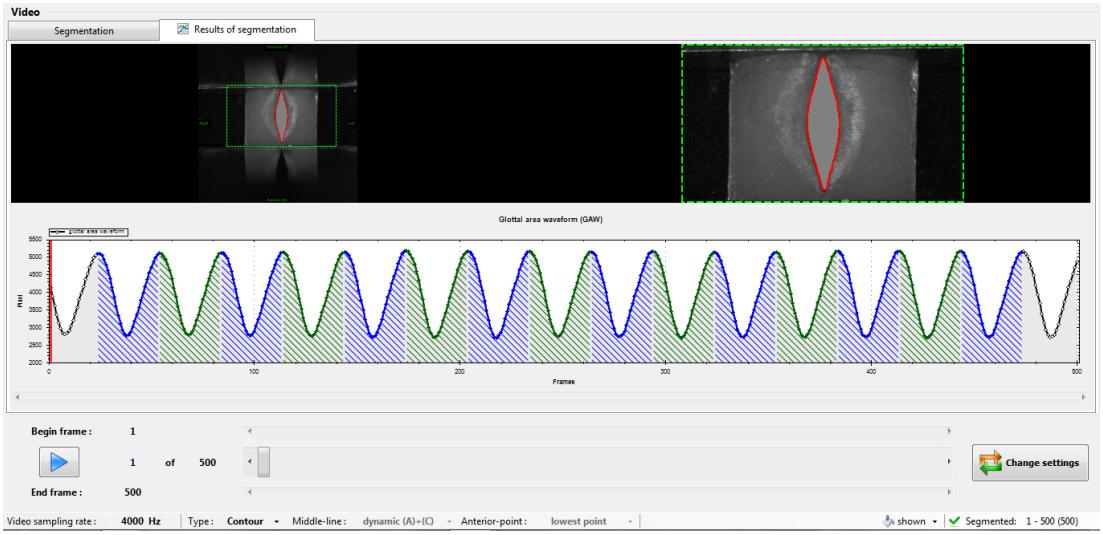
A Figure 21 - Model 113 in 58mm channel configuration in M2 during the setting of the maximum glottal gap with seatpoints with GAT.



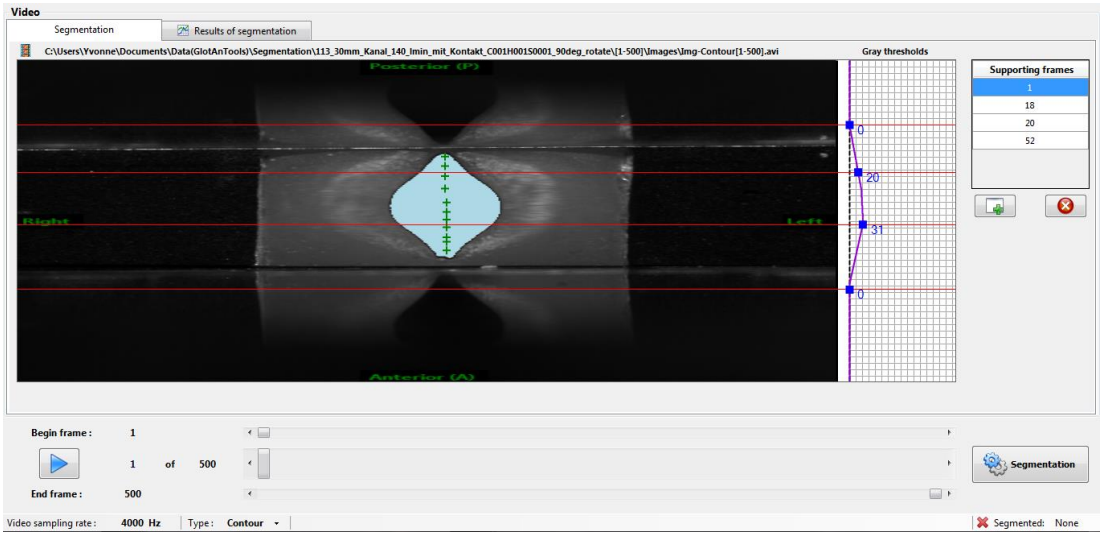
A Figure 22 - Segmentation of the model 113 in 58mm channel configuration M2 with GAT.



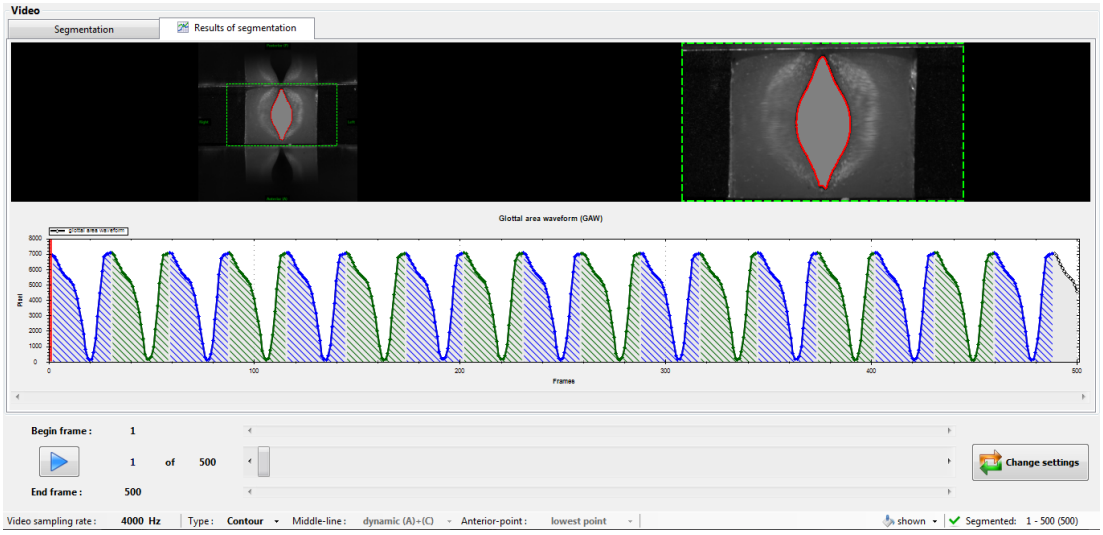
A Figure 23 - Model 113 in 58mm channel configuration in M1 during the setting of the maximum glottal gap with seatpoints with GAT.



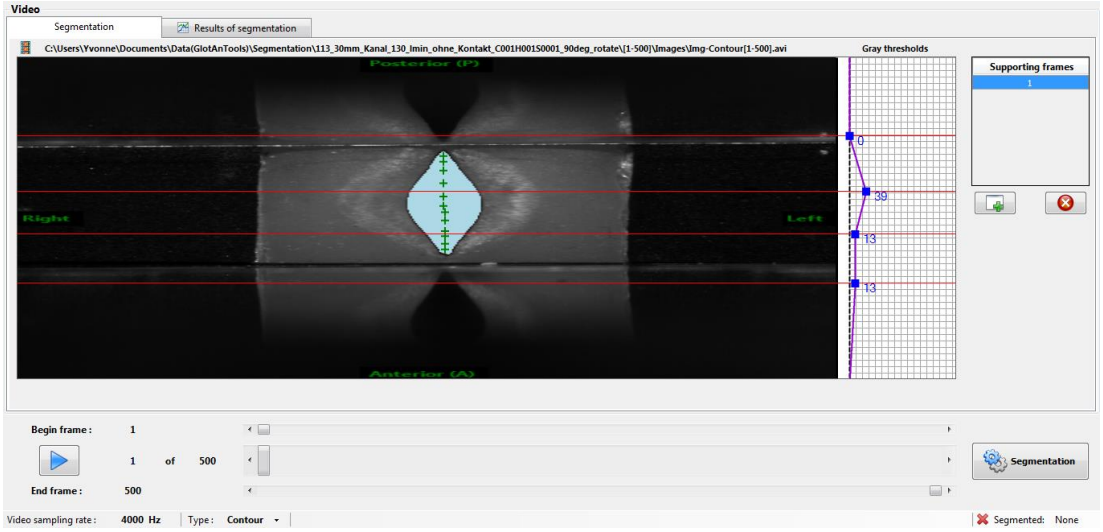
A Figure 24 - Segmentation of the model 113 in 58mm channel configuration M1 with GAT.



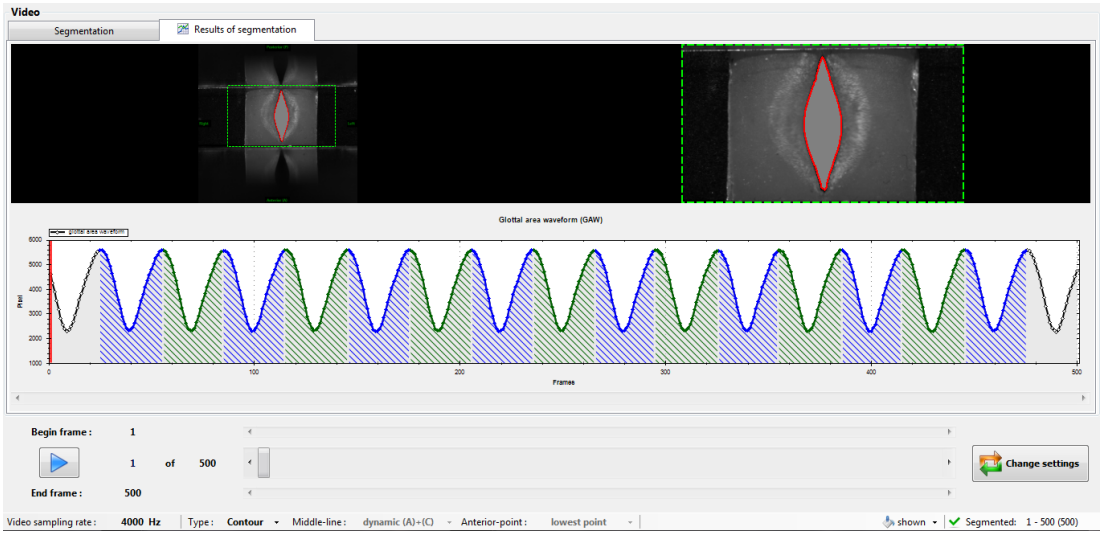
A Figure 25 - Model 113 in 78mm channel configuration in M2 during the setting of the maximum glottal gap with seatpoints with GAT.



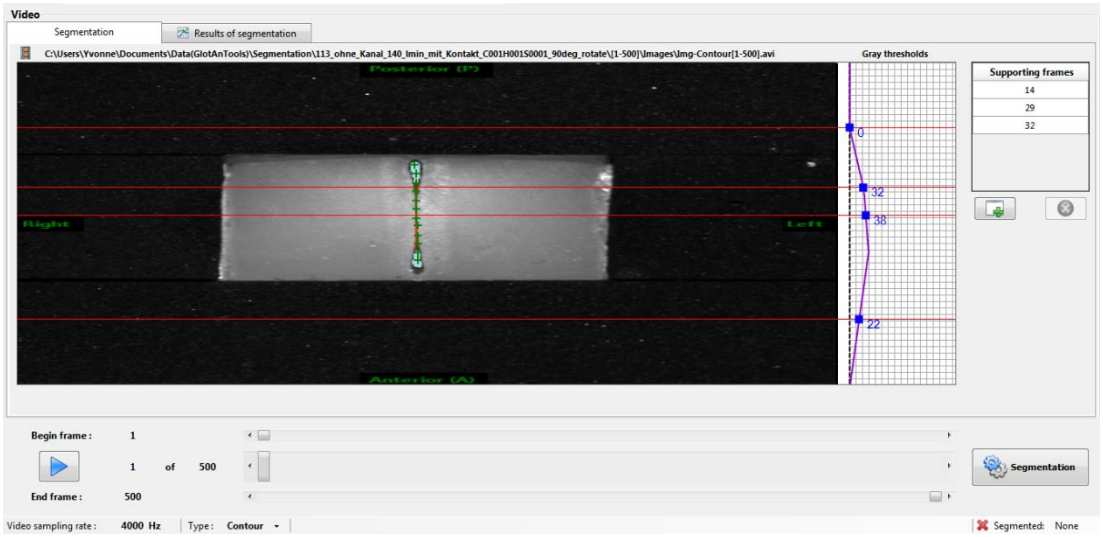
A Figure 26 - Segmentation of the model 113 in 78mm channel configuration M2 with GAT.



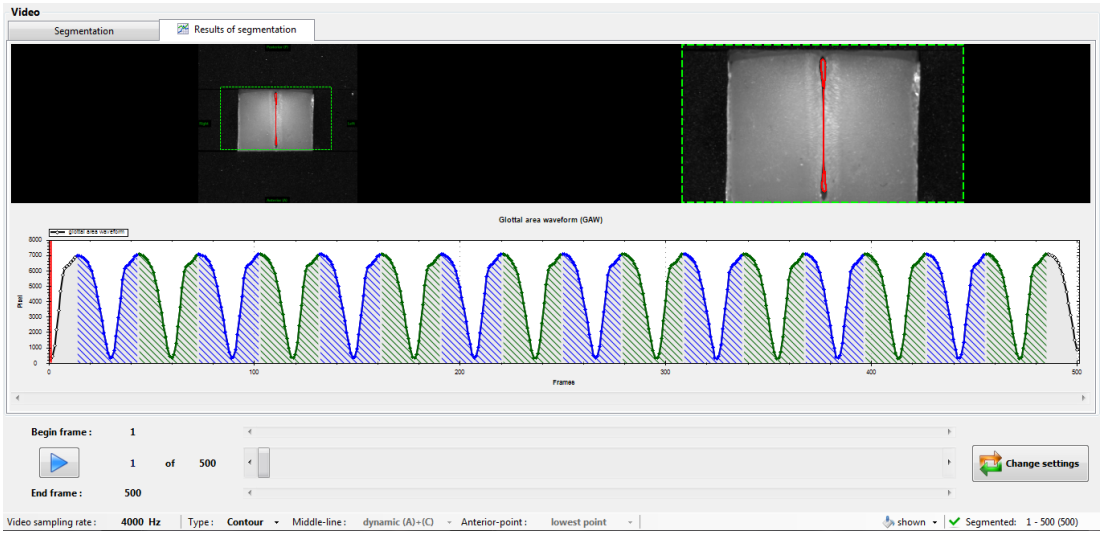
A Figure 27 - Model 113 in 78mm channel configuration in M1 during the setting of the maximum glottal gap with seatpoints with GAT.



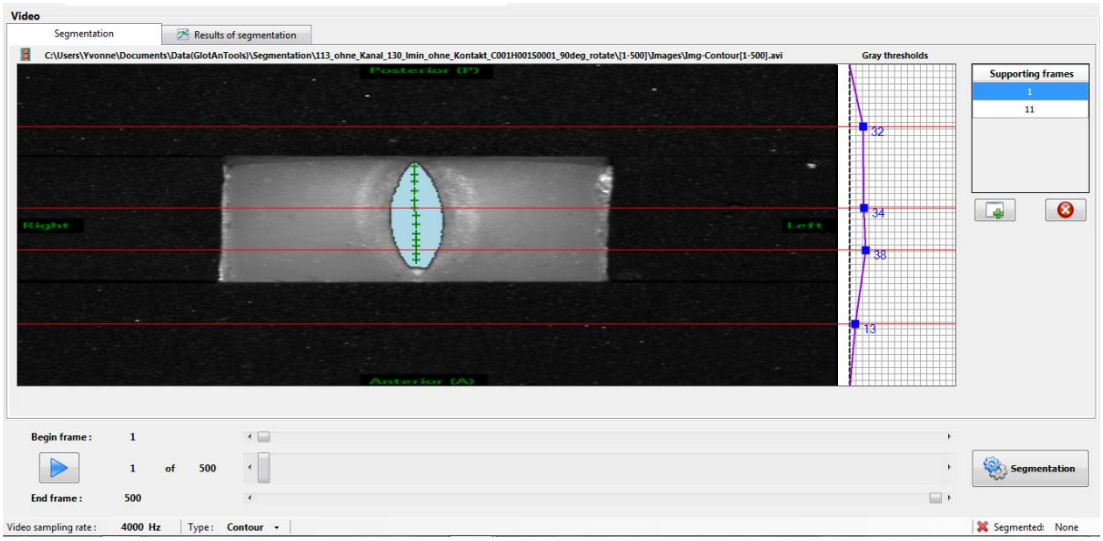
A Figure 28 - Segmentation of the model 113 in 78mm channel configuration M1 with GAT.



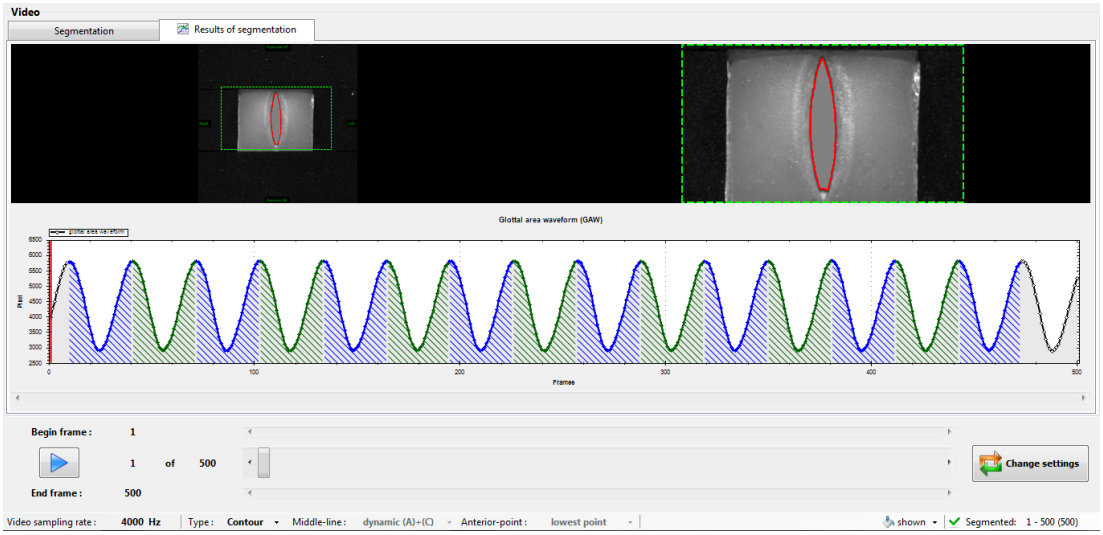
A Figure 29 - Model 113 in no channel channel configuration in M2 during the setting of the maximum glottal gap with seatpoints with GAT.



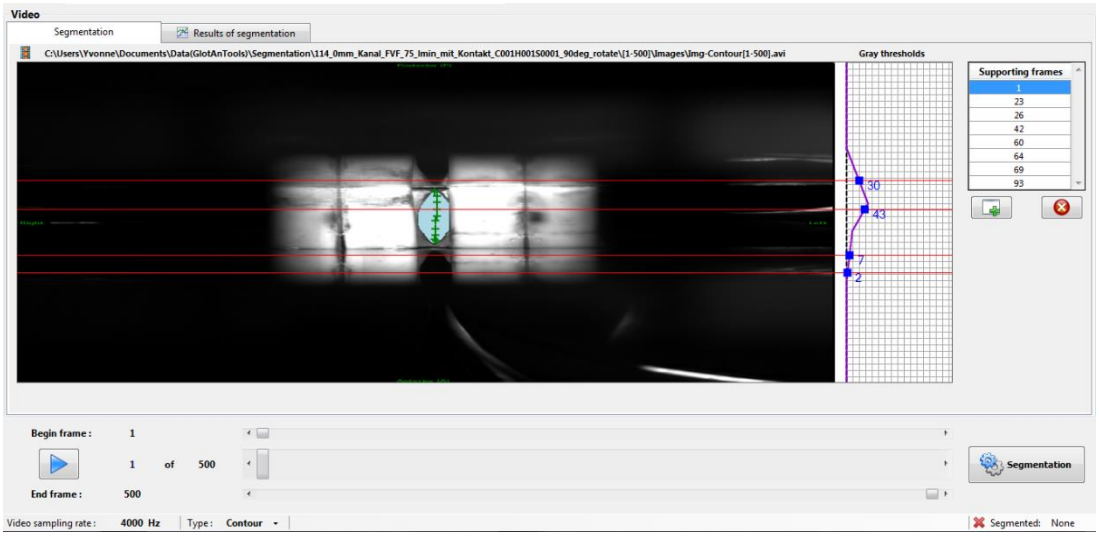
A Figure 30 - Segmentation of the model 113 in no channel channel configuration M2 with GAT.



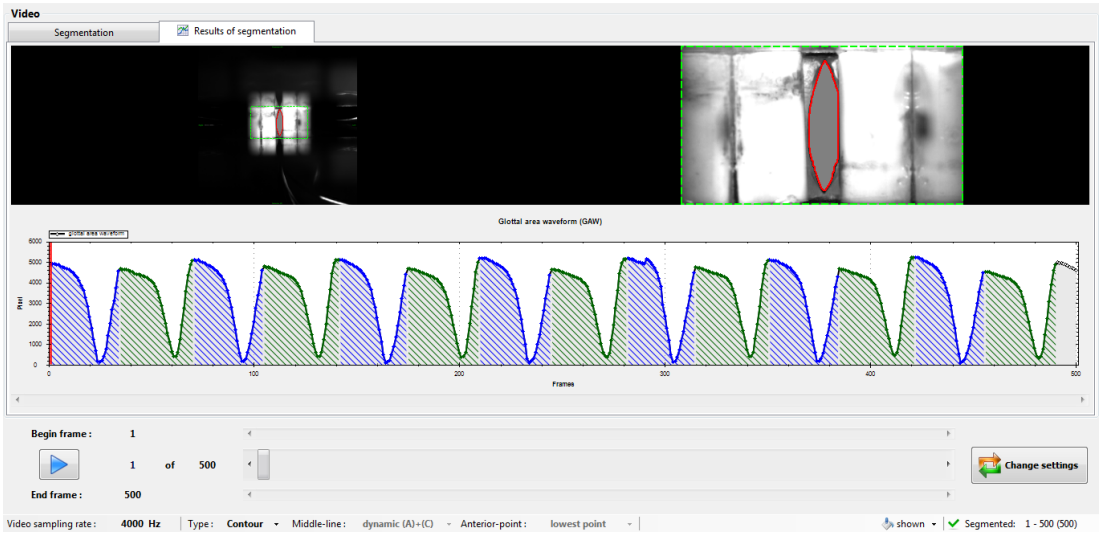
A Figure 31 - Model 113 in no channel channel configuration in M1 during the setting of the maximum glottal gap with seatpoints with GAT.



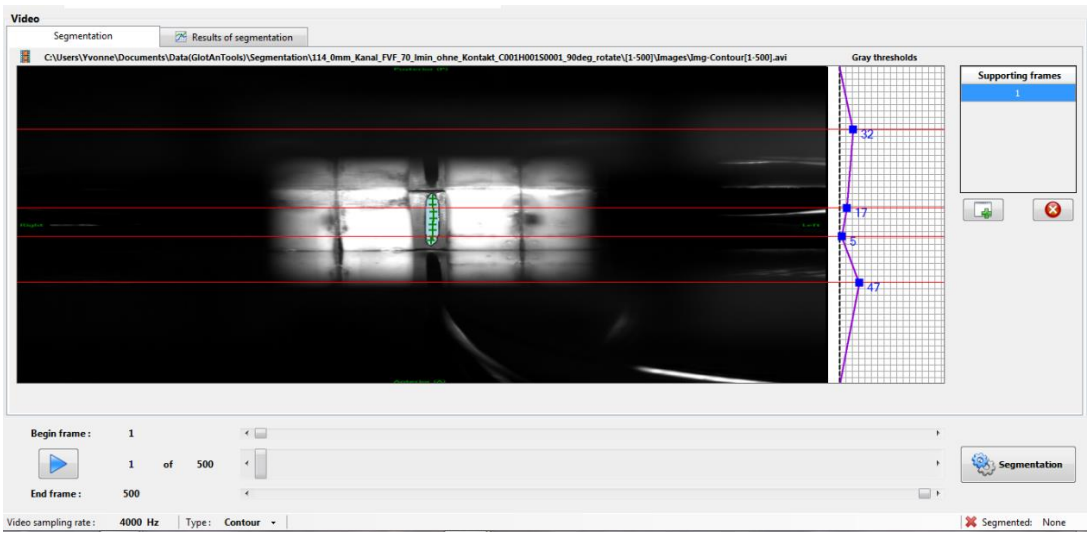
A Figure 32 - Segmentation of the model 113 in no channel channel configuration M1 with GAT.



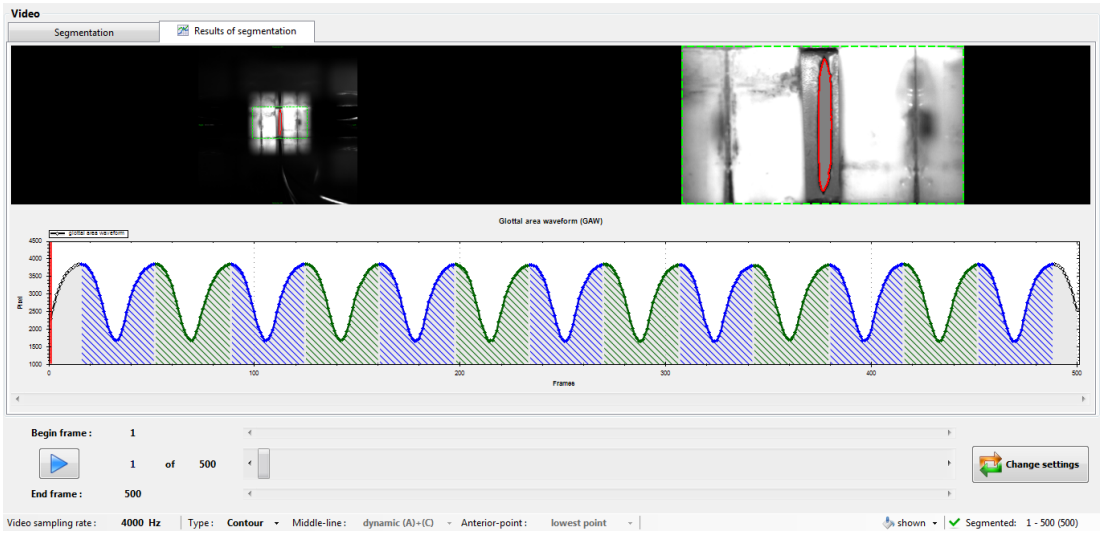
A Figure 33 - Model 114 in 18mm with FVF channel configuration in M2 during the setting of the maximum glottal gap with seatpoints with GAT.



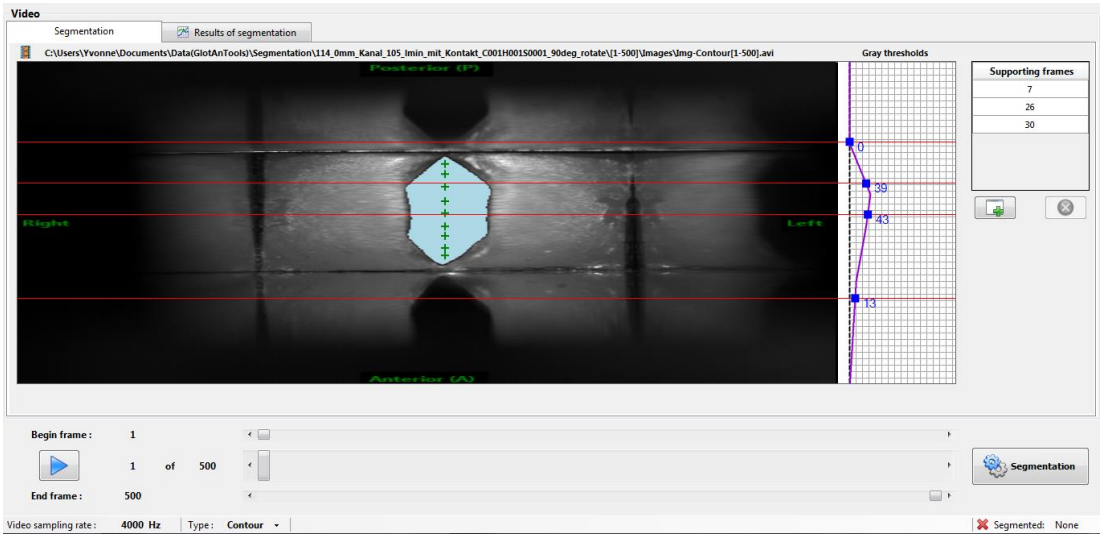
A Figure 34 - Segmentation of the model 114 in 18mm with FVF channel configuration M2 with GAT.



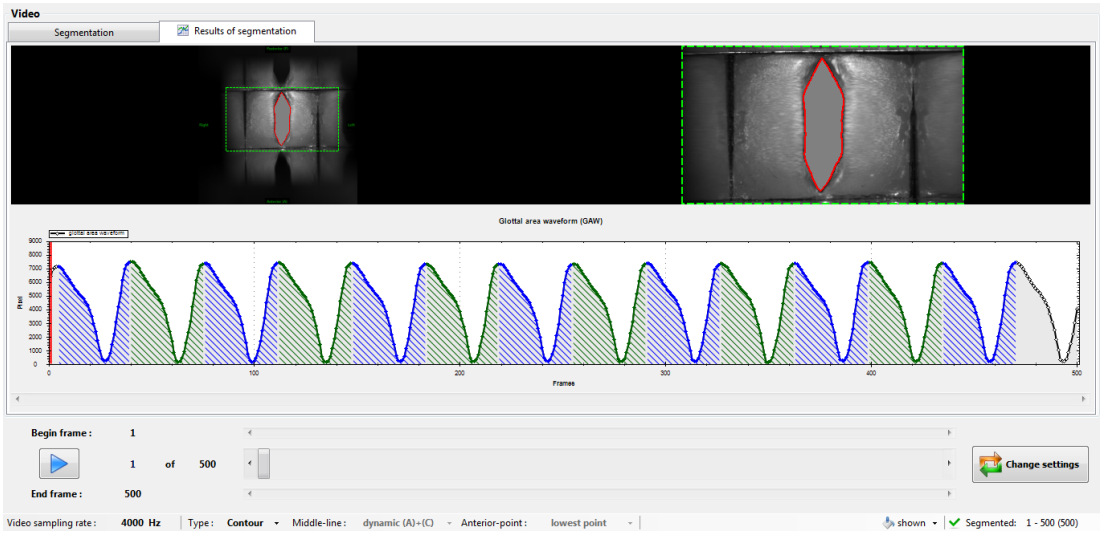
A Figure 35 - Model 114 in 18mm with FVF channel configuration in M1 during the setting of the maximum glottal gap with seatpoints with GAT.



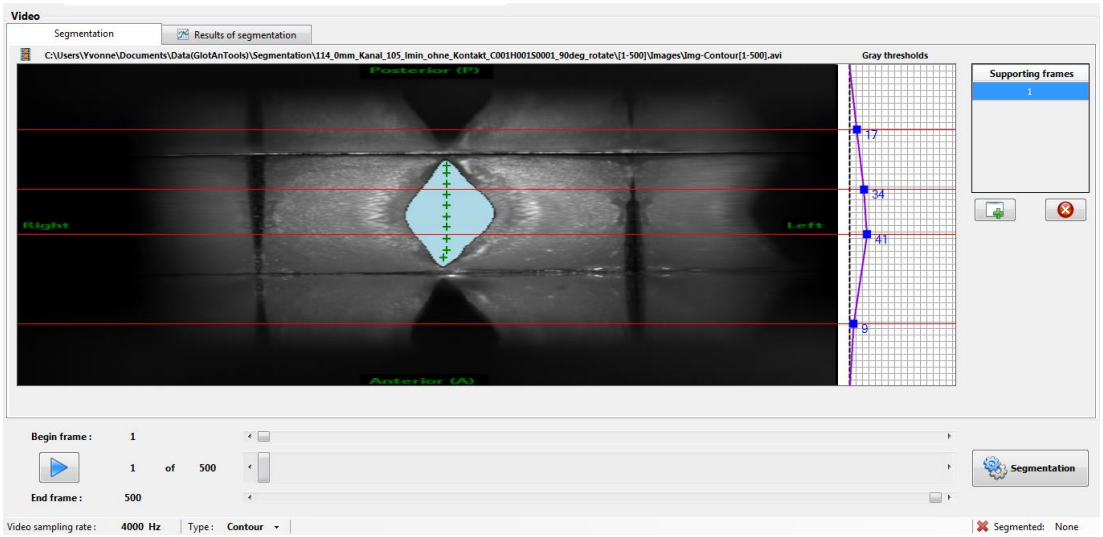
A Figure 36 - Segmentation of the model 114 in 18mm with FVF channel configuration M1 with GAT.



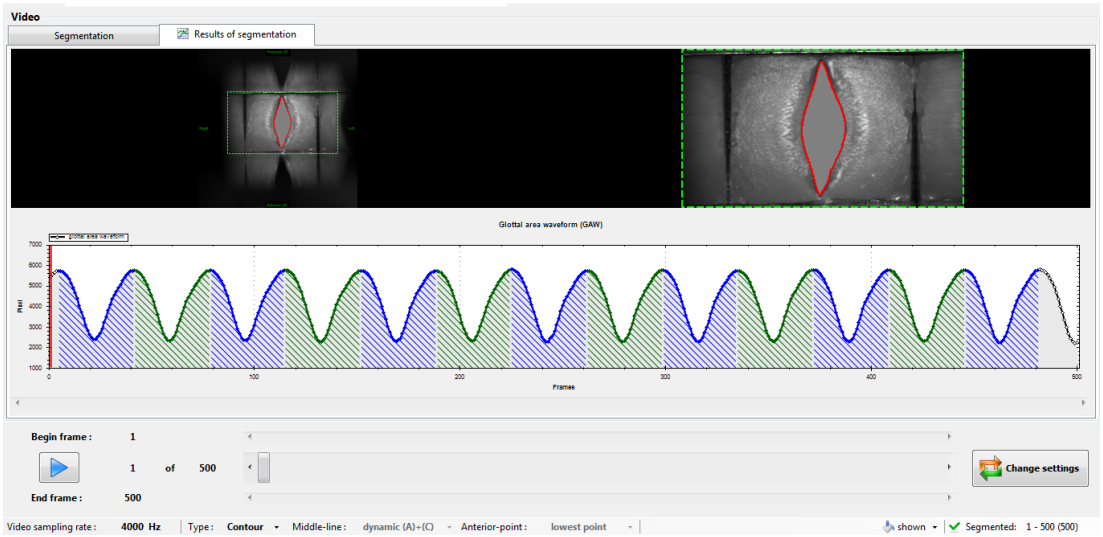
A Figure 37 - Model 114 in 18mm channel configuration in M2 during the setting of the maximum glottal gap with seatpoints with GAT.



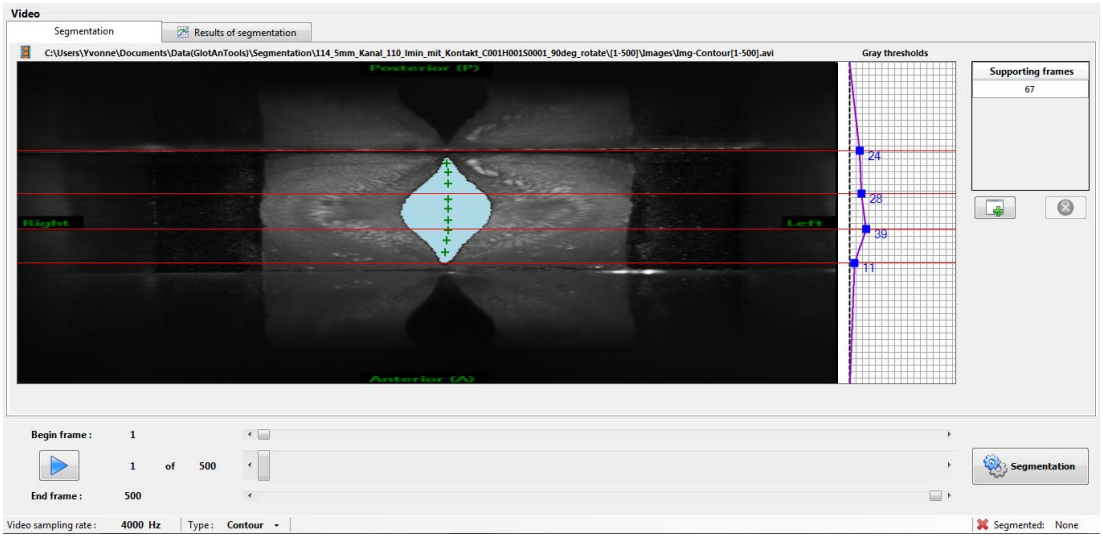
A Figure 38 - Segmentation of the model 114 in 18mm channel configuration M2 with GAT.



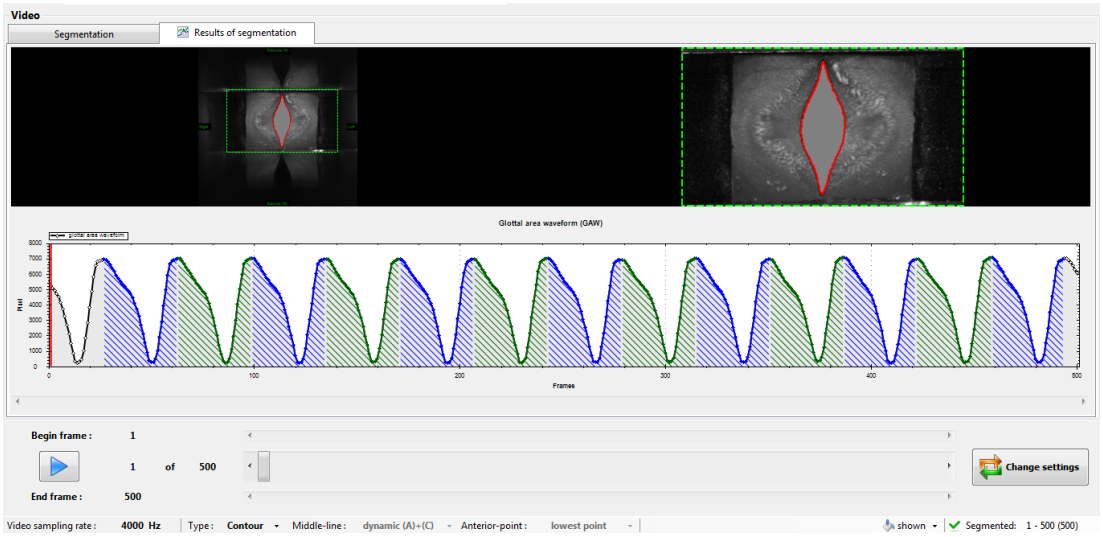
A Figure 39 - Model 114 in 18mm channel configuration in M1 during the setting of the maximum glottal gap with seatpoints with GAT.



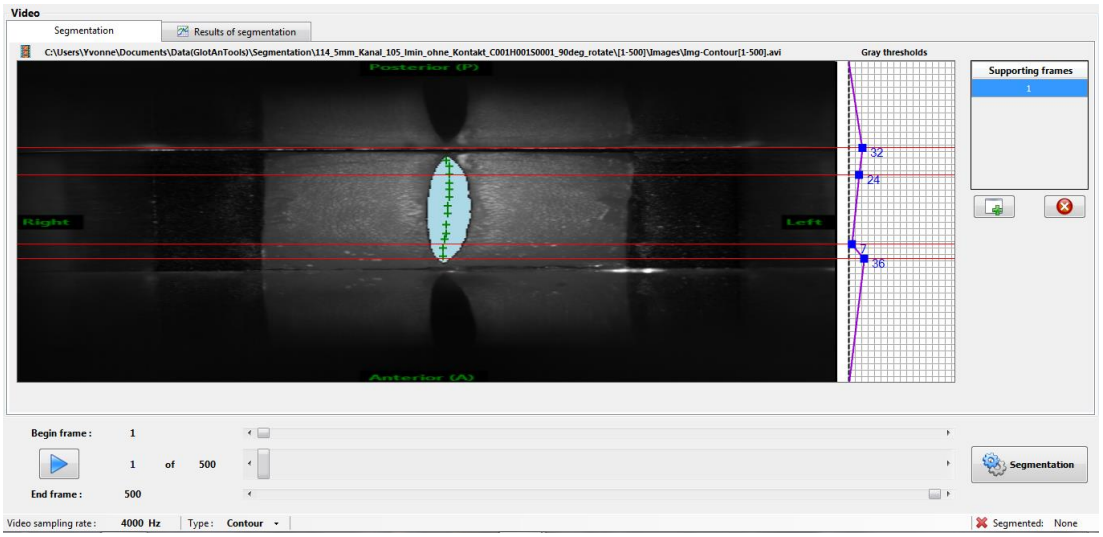
A Figure 40 - Segmentation of the model 114 in 18mm channel configuration M1 with GAT.



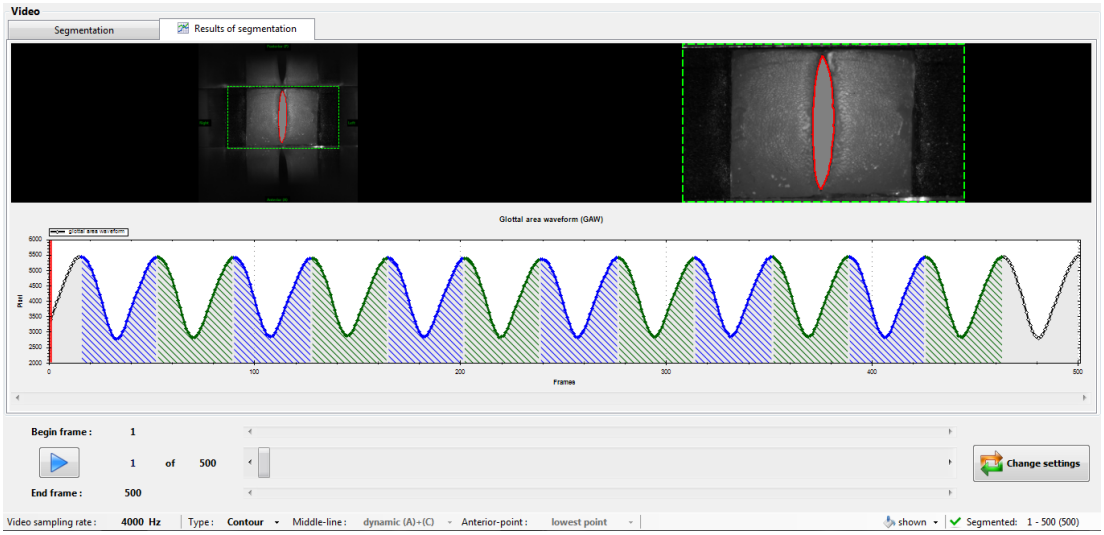
A Figure 41 - Model 114 in 28mm channel configuration in M2 during the setting of the maximum glottal gap with seatpoints with GAT.



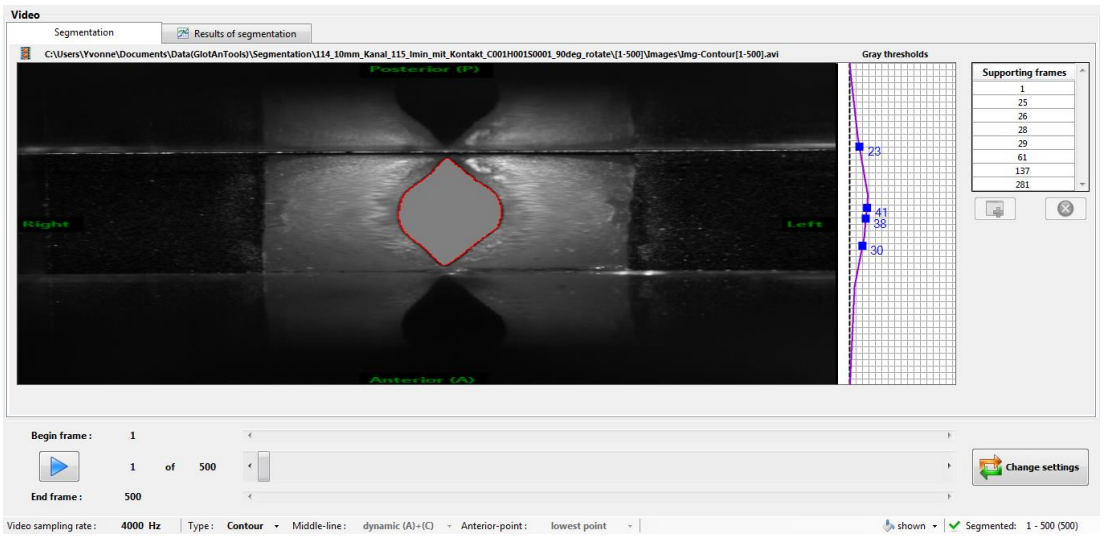
A Figure 42 - Segmentation of the model 114 in 28mm channel configuration M2 with GAT.



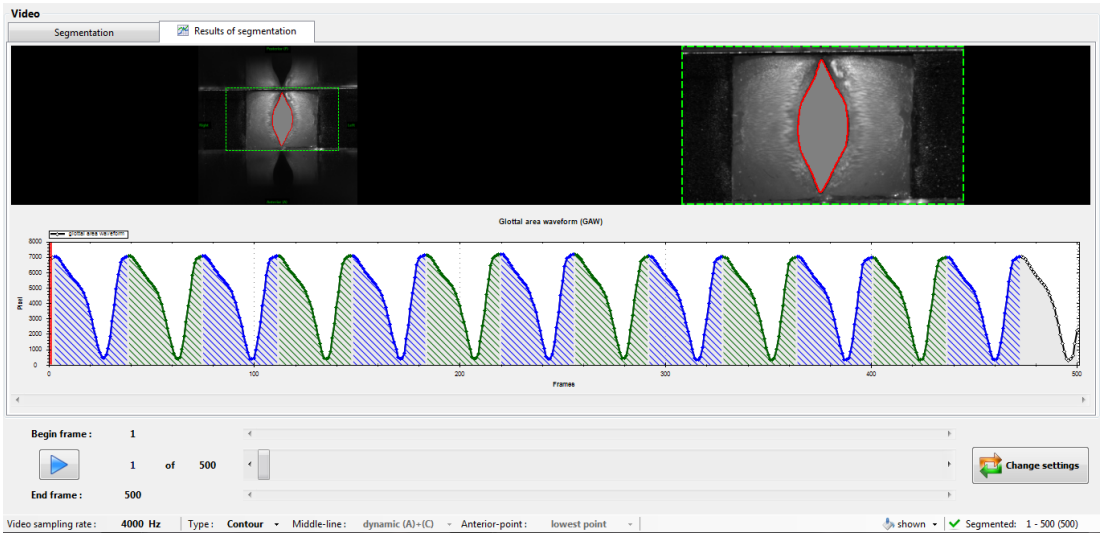
A Figure 43 - Model 114 in 28mm channel configuration in M1 during the setting of the maximum glottal gap with seatpoints with GAT.



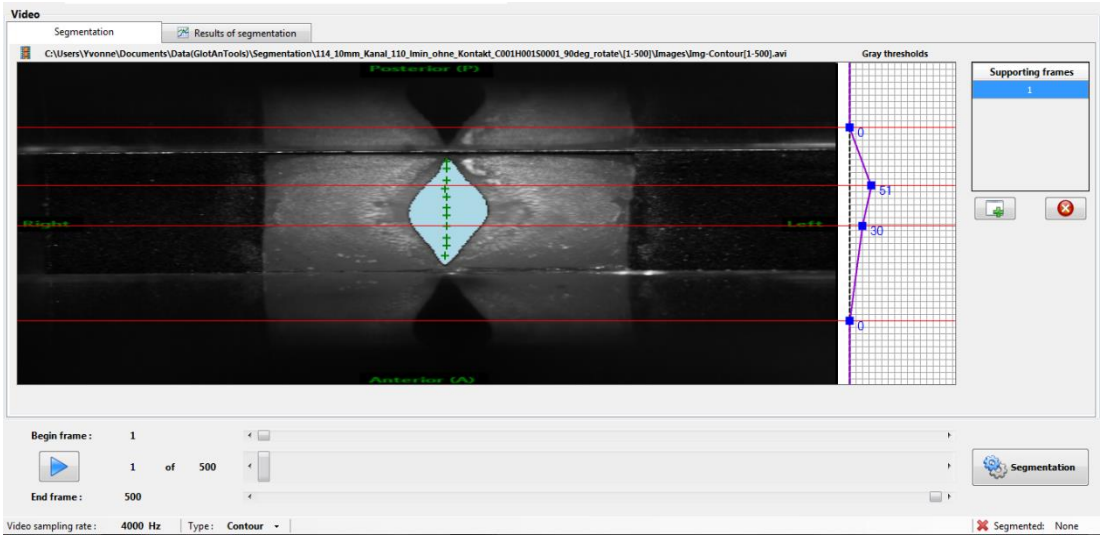
A Figure 44 - Segmentation of the model 114 in 28mm channel configuration M1 with GAT.



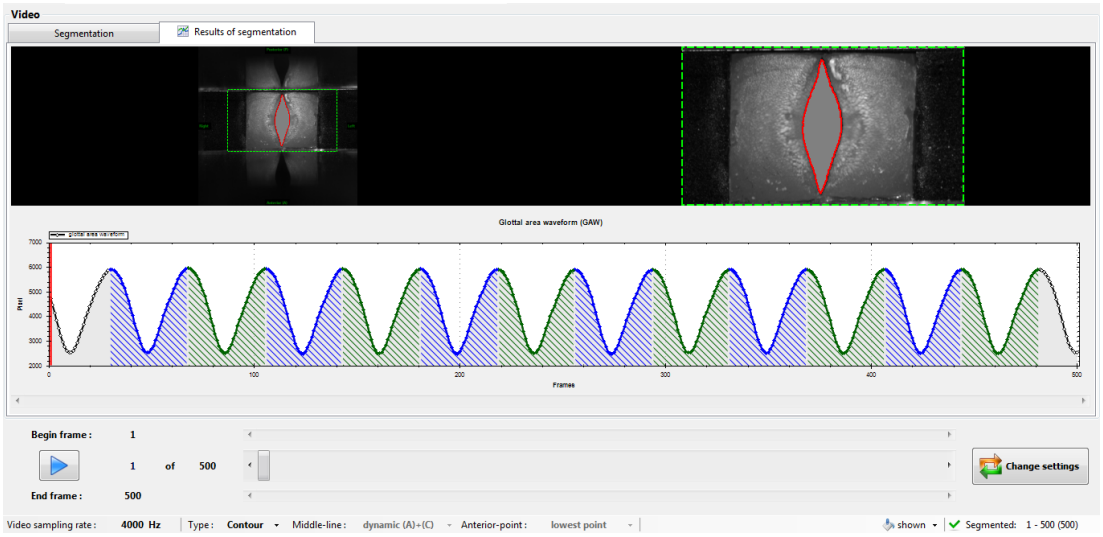
A Figure 45 - Model 114 in 38mm channel configuration in M2 during the setting of the maximum glottal gap with seatpoints with GAT.



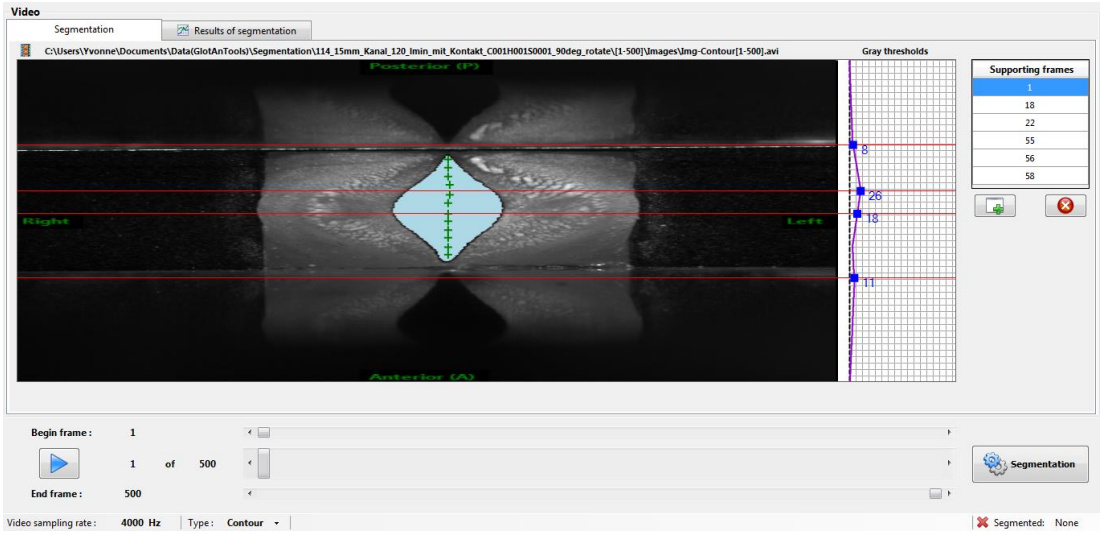
A Figure 46 - Segmentation of the model 114 in 38mm channel configuration M2 with GAT.



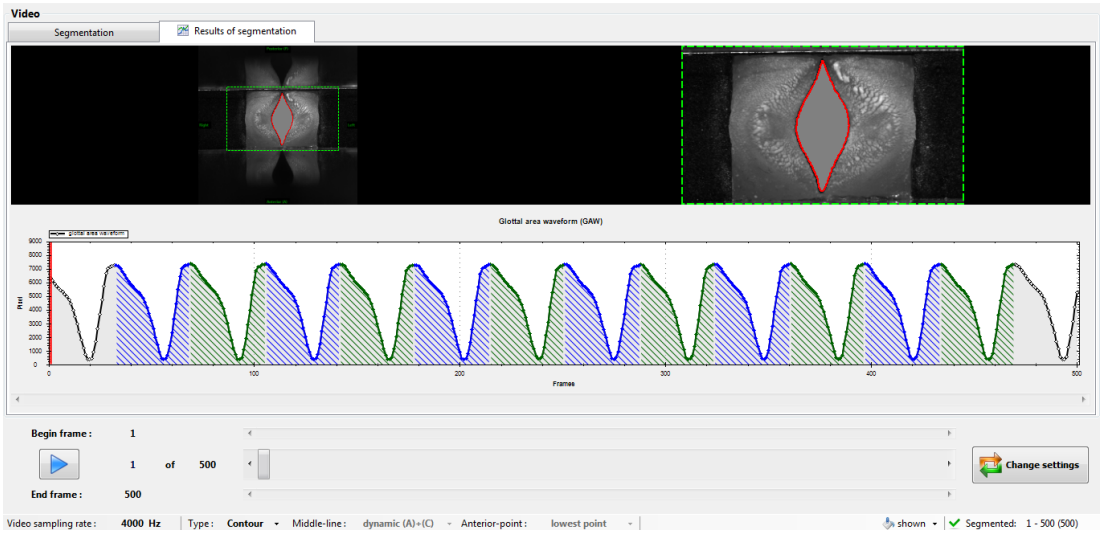
A Figure 47 - Model 114 in 38mm channel configuration in M1 during the setting of the maximum glottal gap with seatpoints with GAT.



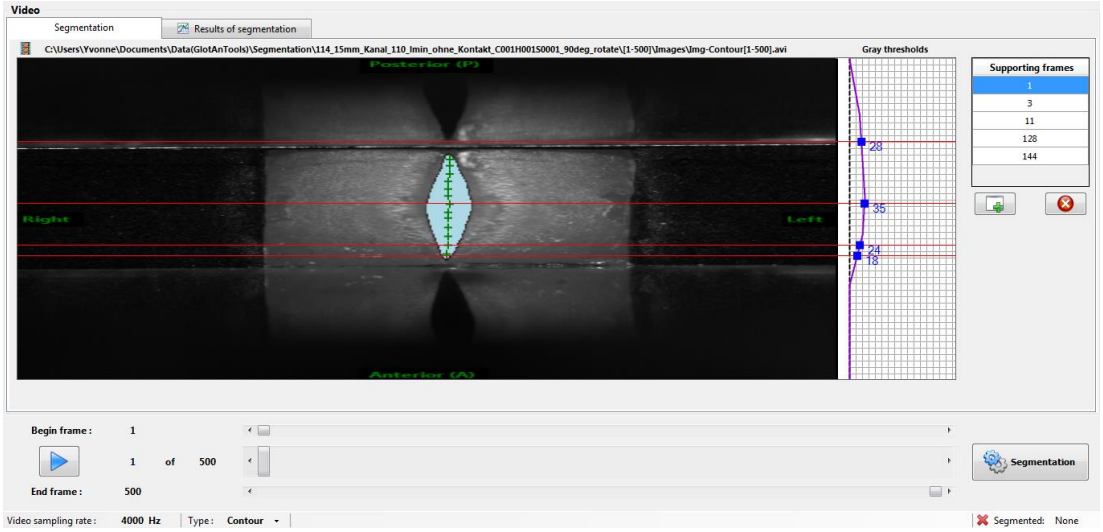
A Figure 48 - Segmentation of the model 114 in 38mm channel configuration M1 with GAT.



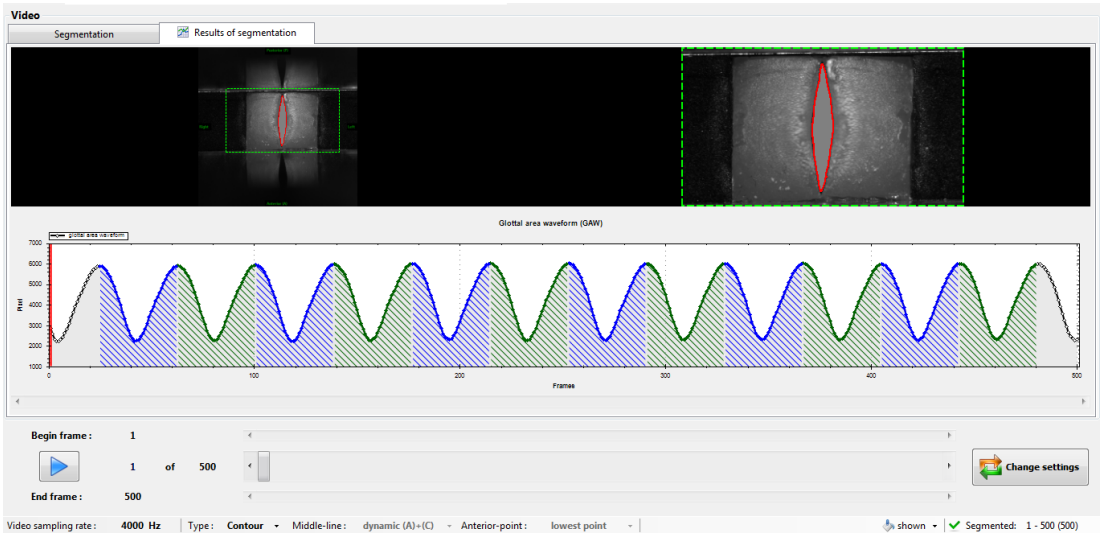
A Figure 49 - Model 114 in 48mm channel configuration in M2 during the setting of the maximum glottal gap with seatpoints with GAT.



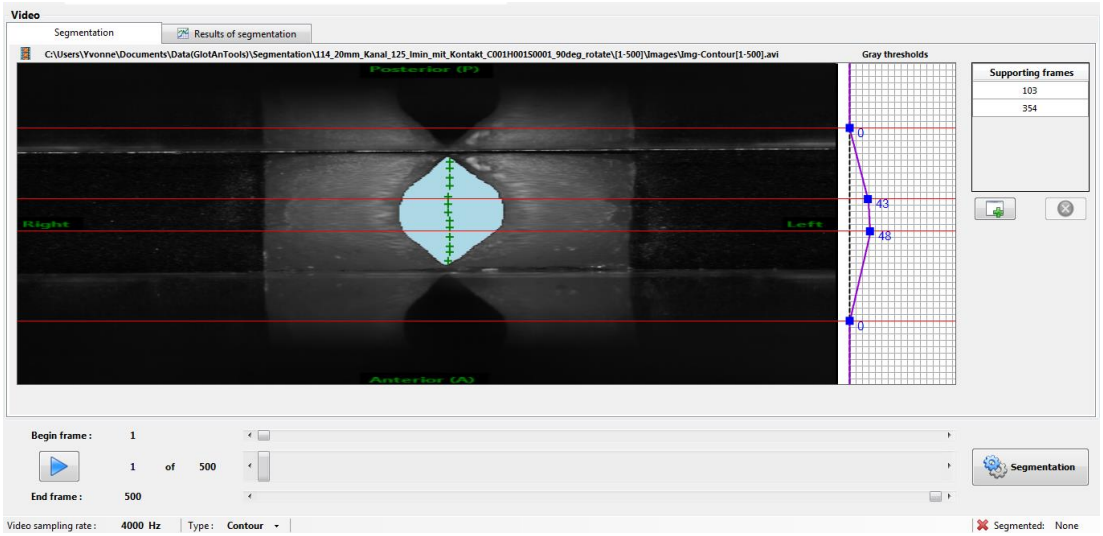
A Figure 50 - Segmentation of the model 114 in 48mm channel configuration M2 with GAT.



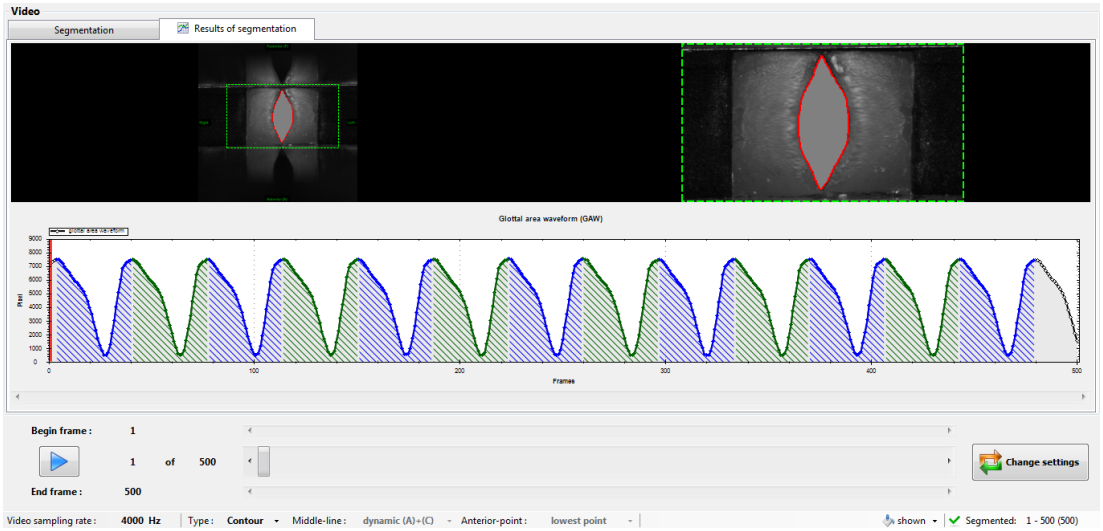
A Figure 51 - Model 114 in 48mm channel configuration in M1 during the setting of the maximum glottal gap with seatpoints with GAT.



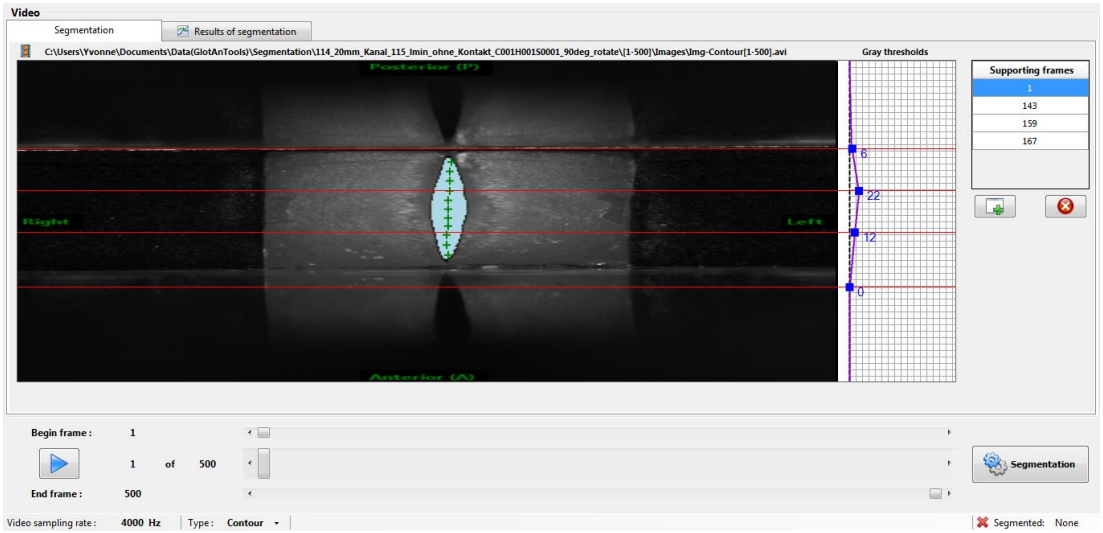
A Figure 52 - Segmentation of the model 114 in 48mm channel configuration M1 with GAT.



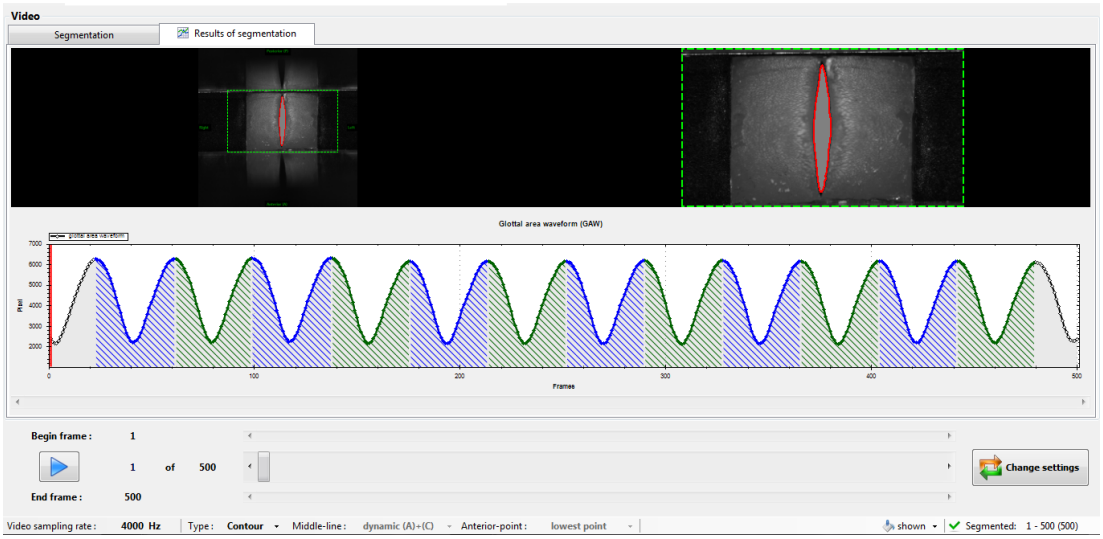
A Figure 53 - Model 114 in 58mm channel configuration in M2 during the setting of the maximum glottal gap with seatpoints with GAT.



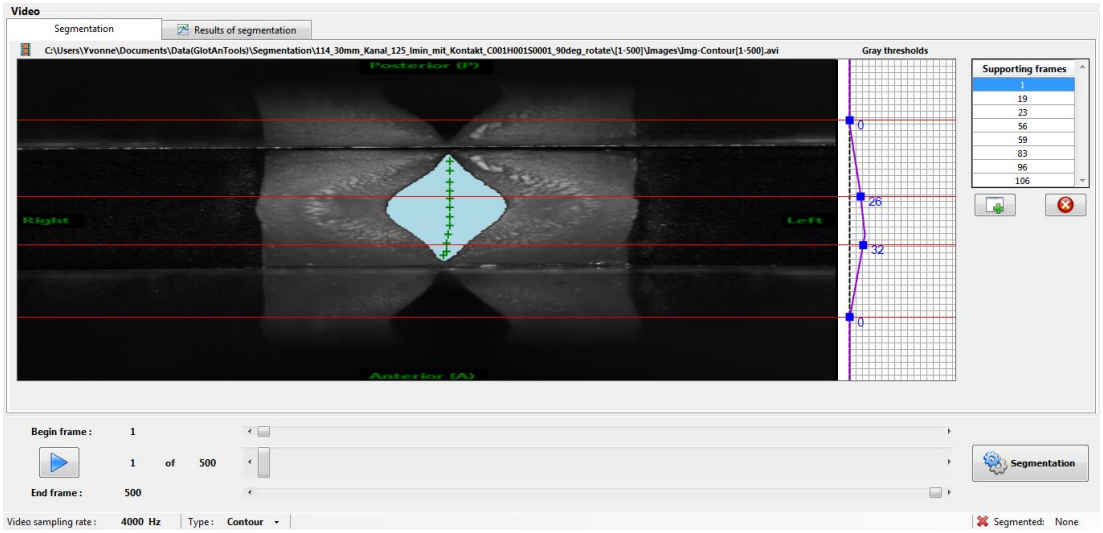
A Figure 54 - Segmentation of the model 114 in 58mm channel configuration M2 with GAT.



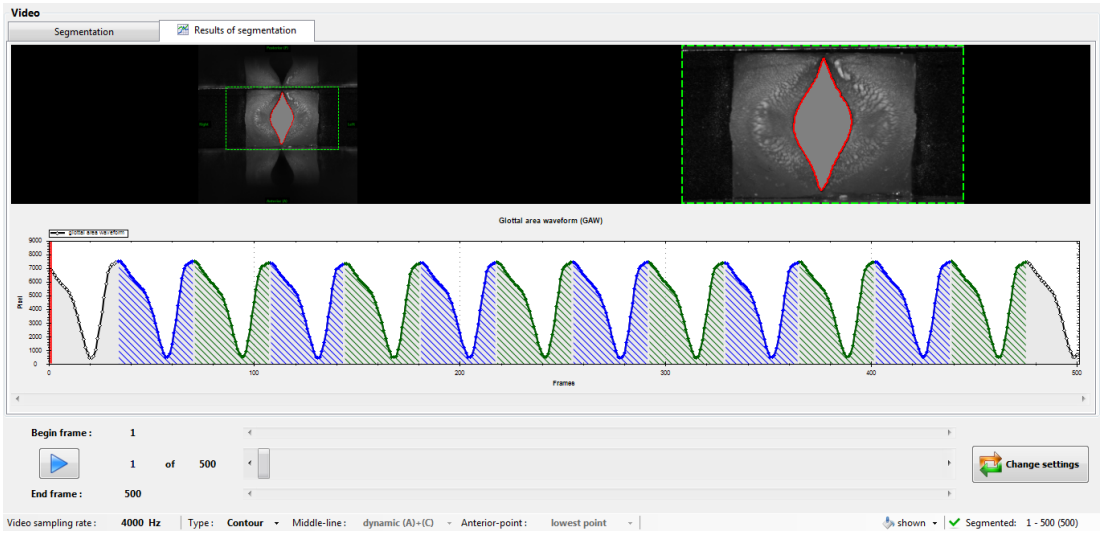
A Figure 55 - Model 114 in 58mm channel configuration in M1 during the setting of the maximum glottal gap with seatpoints with GAT.



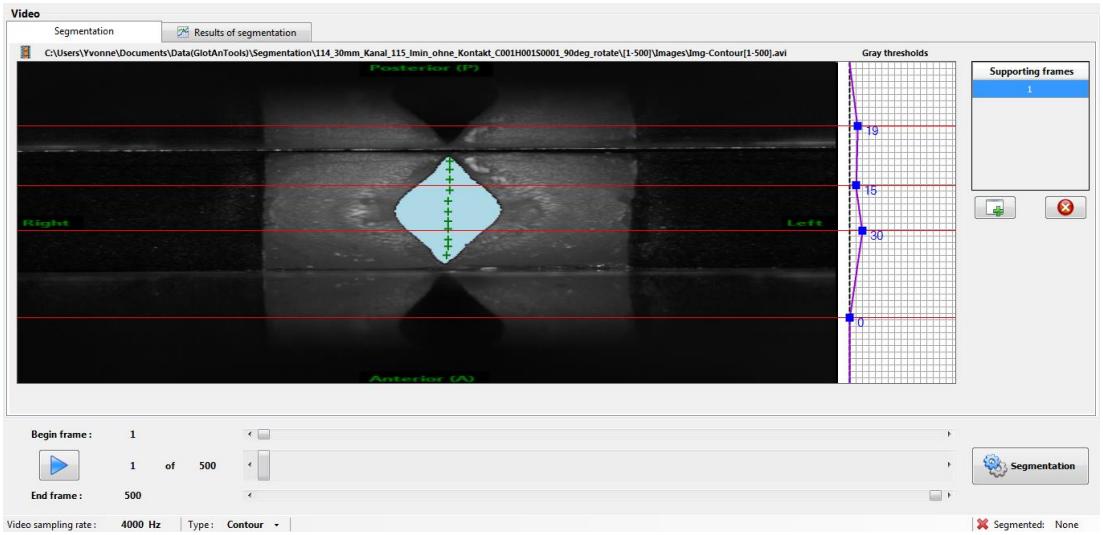
A Figure 56 - Segmentation of the model 114 in 58mm channel configuration M1 with GAT.



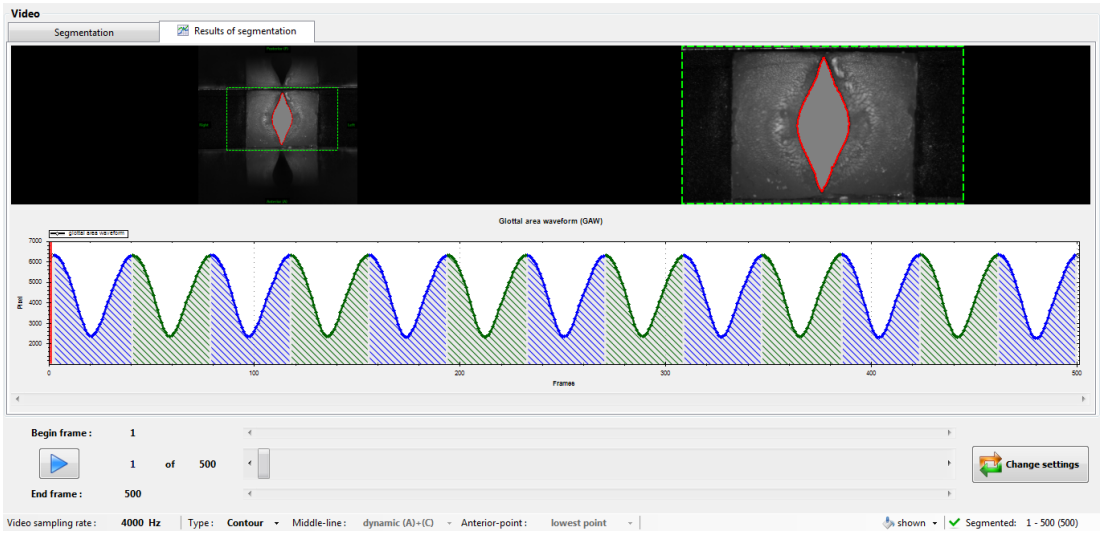
A Figure 57 - Model 114 in 78mm channel configuration in M2 during the setting of the maximum glottal gap with seatpoints with GAT.



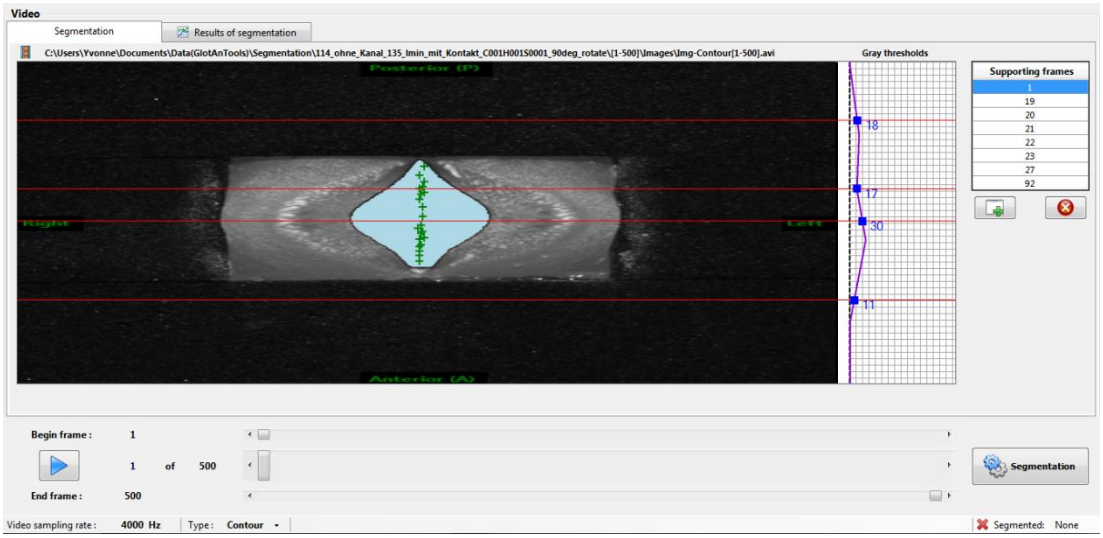
A Figure 58 - Segmentation of the model 114 in 78mm channel configuration M2 with GAT.



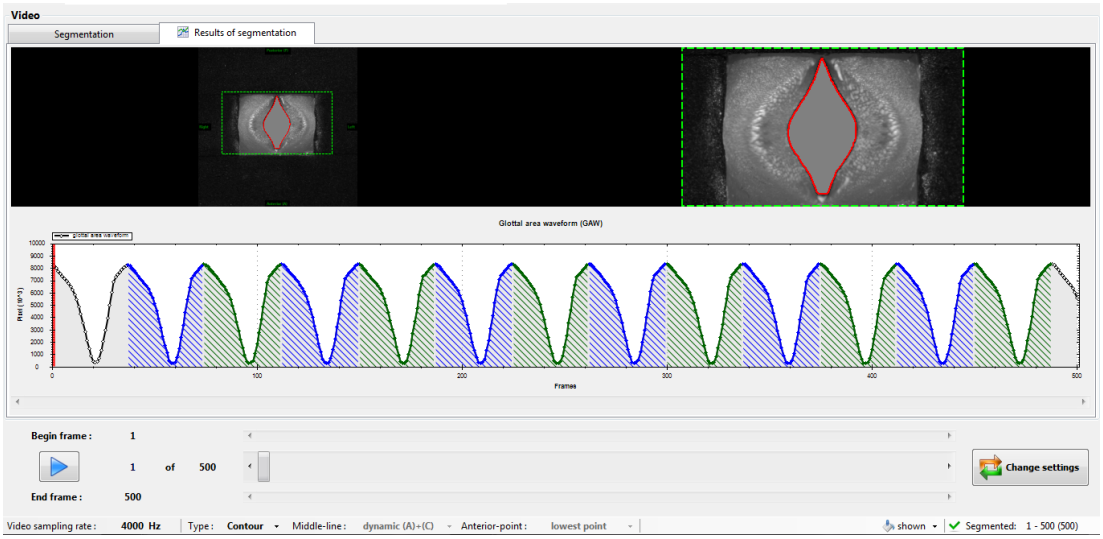
A Figure 59 - Model 114 in 78mm channel configuration in M1 during the setting of the maximum glottal gap with seatpoints with GAT.



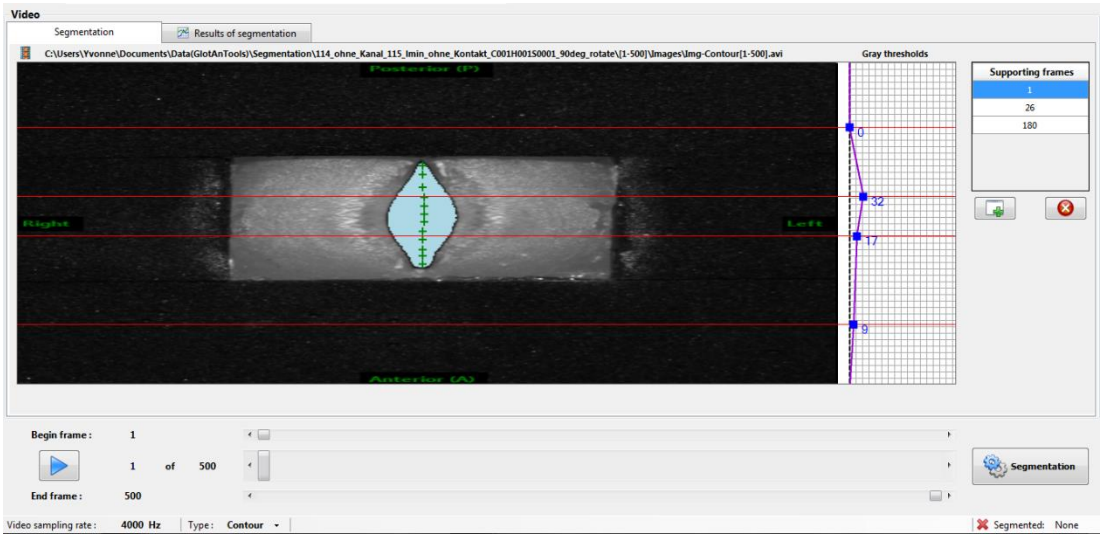
A Figure 60 - Segmentation of the model 114 in 78mm channel configuration M1 with GAT.



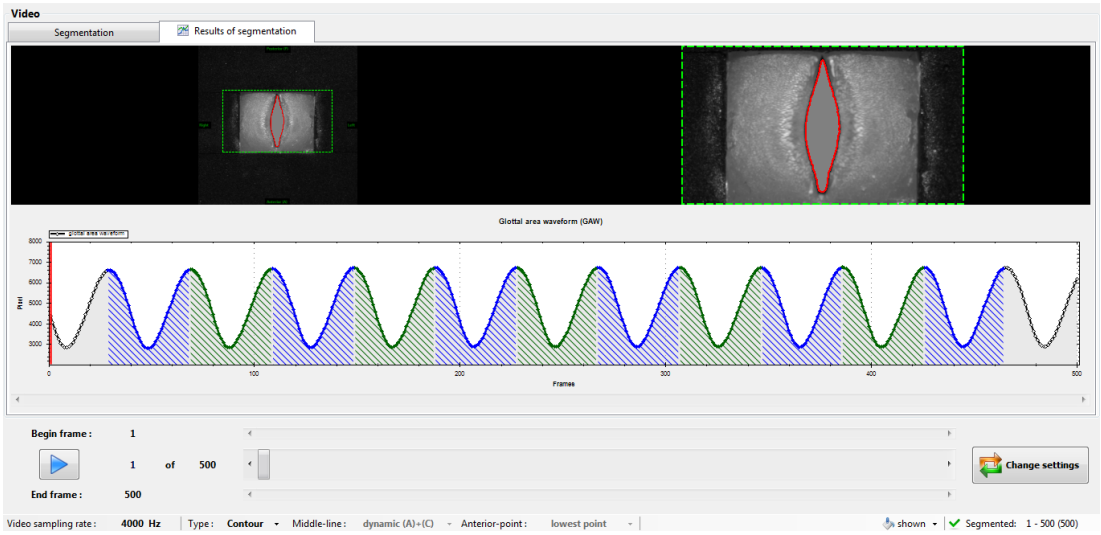
A Figure 61 - Model 114 in no channel channel configuration in M2 during the setting of the maximum glottal gap with seatpoints with GAT.



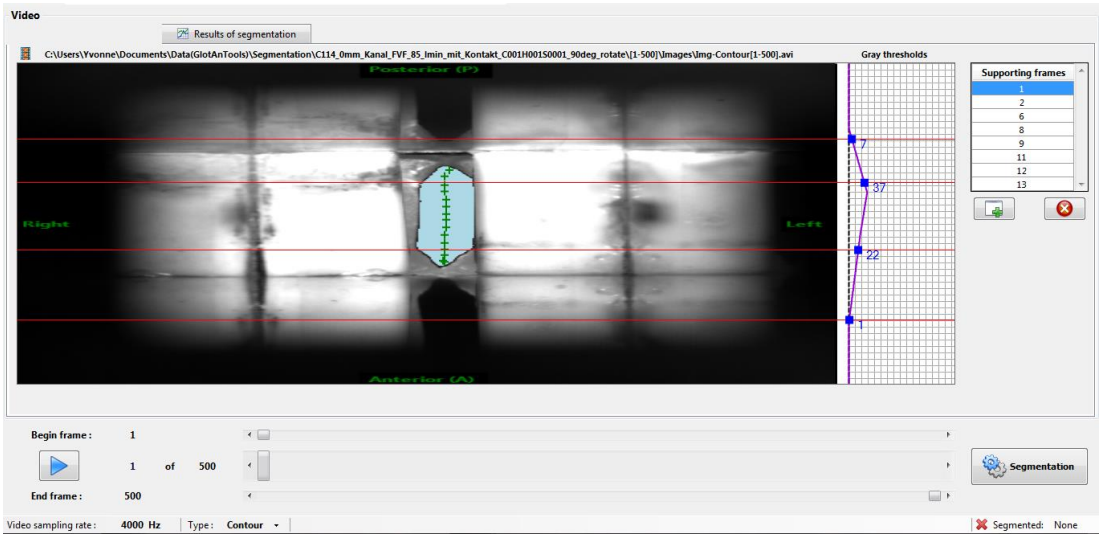
A Figure 62 - Segmentation of the model 114 in no channel channel configuration M2 with GAT.



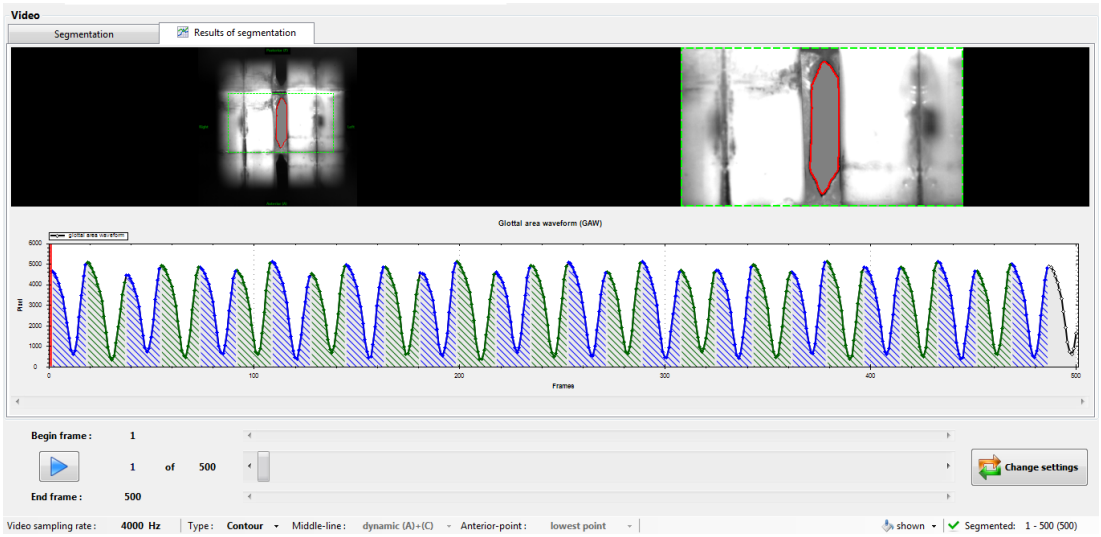
A Figure 63 - Model 114 in no channel channel configuration in M1 during the setting of the maximum glottal gap with seatpoints with GAT.



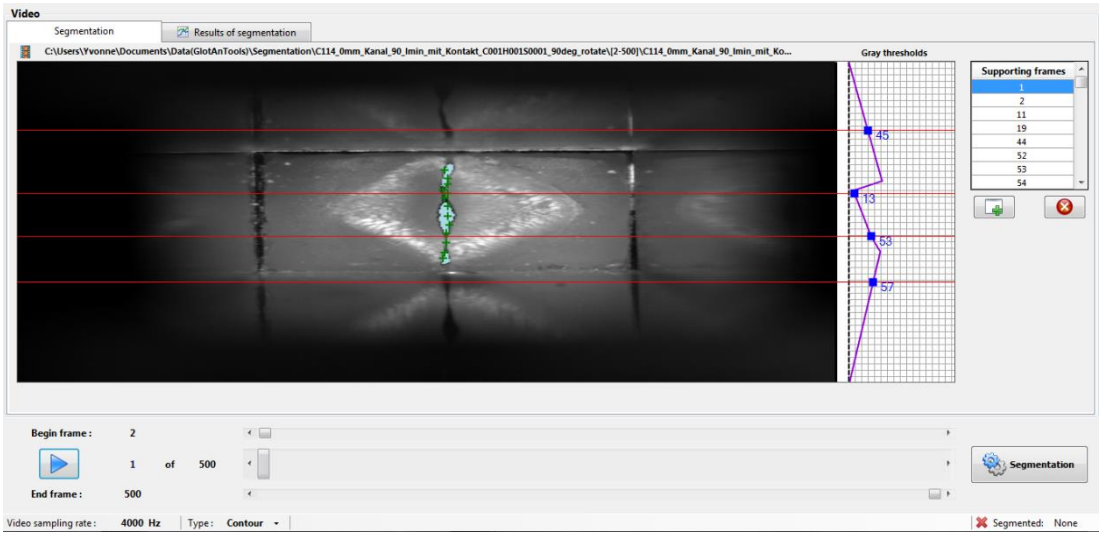
A Figure 64 - Segmentation of the model 114 in no channel channel configuration M1 with GAT.



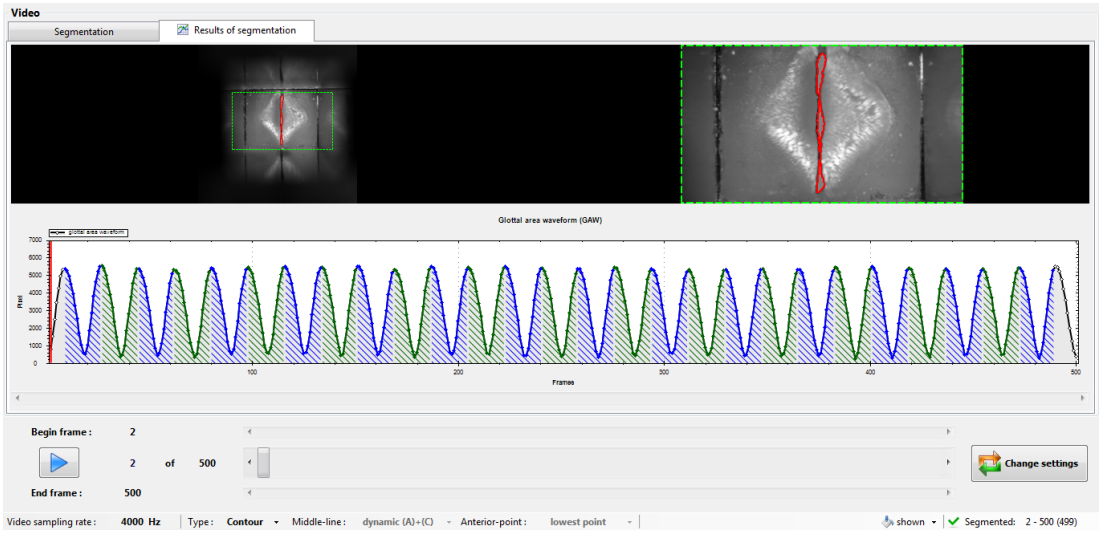
A Figure 65 - Model C114 in 18mm with FVF channel configuration in M2 during the setting of the maximum glottal gap with seatpoints with GAT.



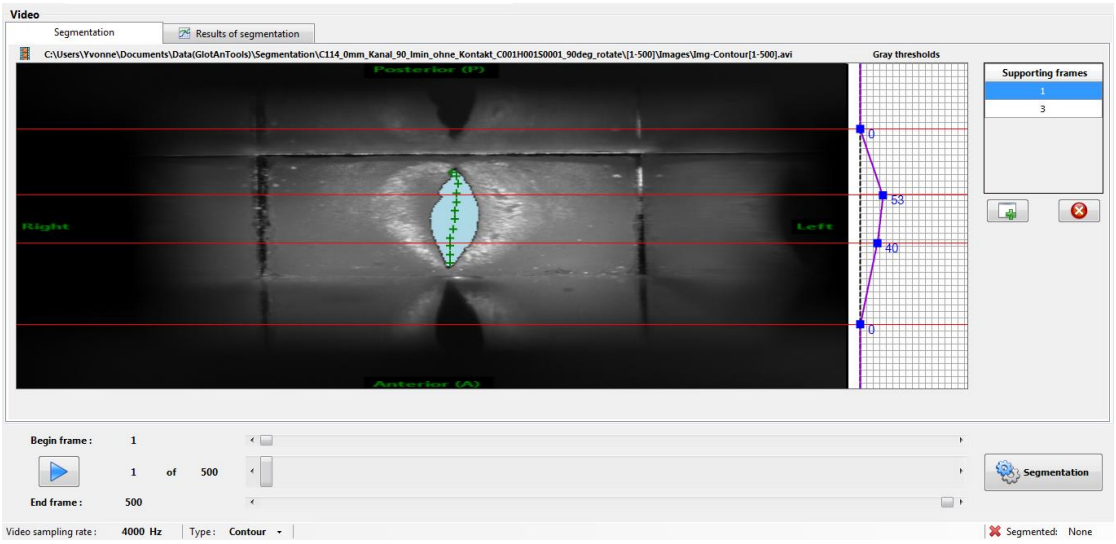
A Figure 66 - Segmentation of the model C114 in 18mm with FVF channel configuration M2 with GAT.



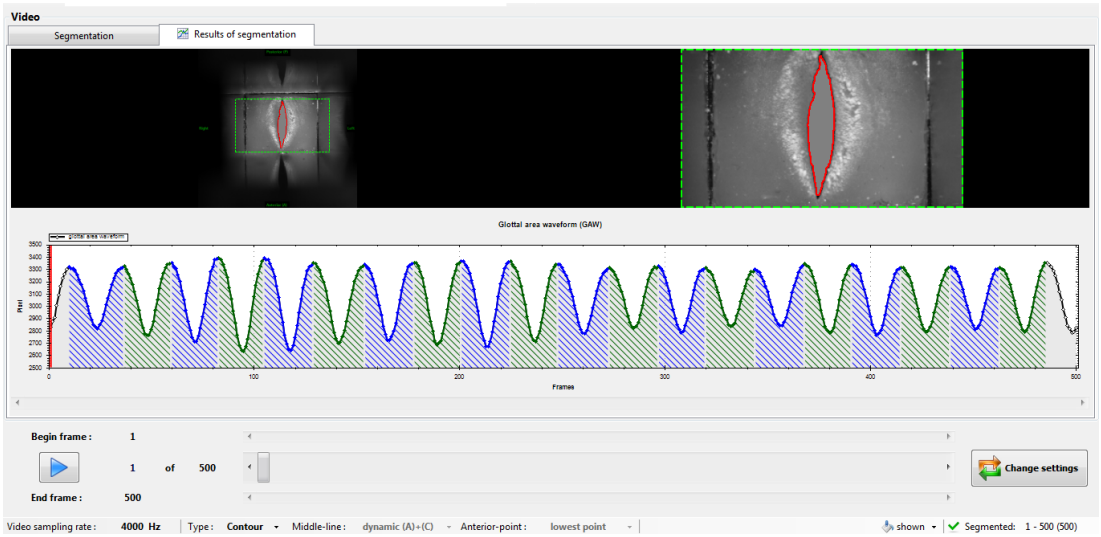
A Figure 67 - Model C114 in 18mm channel configuration in M2 during the setting of the maximum glottal gap with seatpoints with GAT.



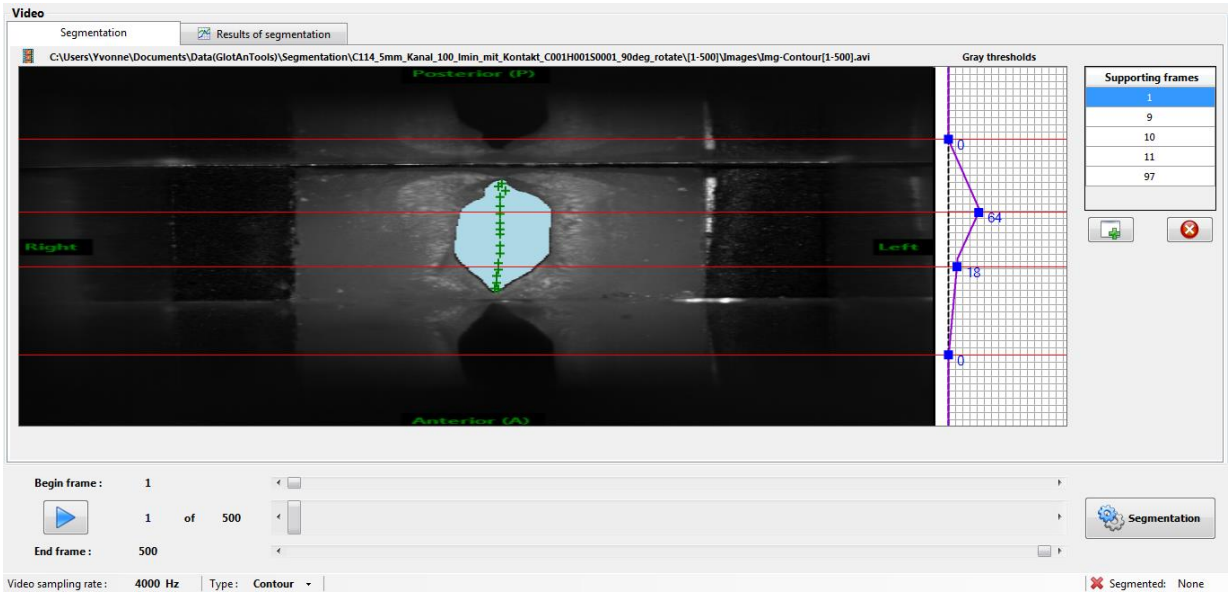
A Figure 68 - Segmentation of the model C114 in 18mm channel configuration M2 with GAT.



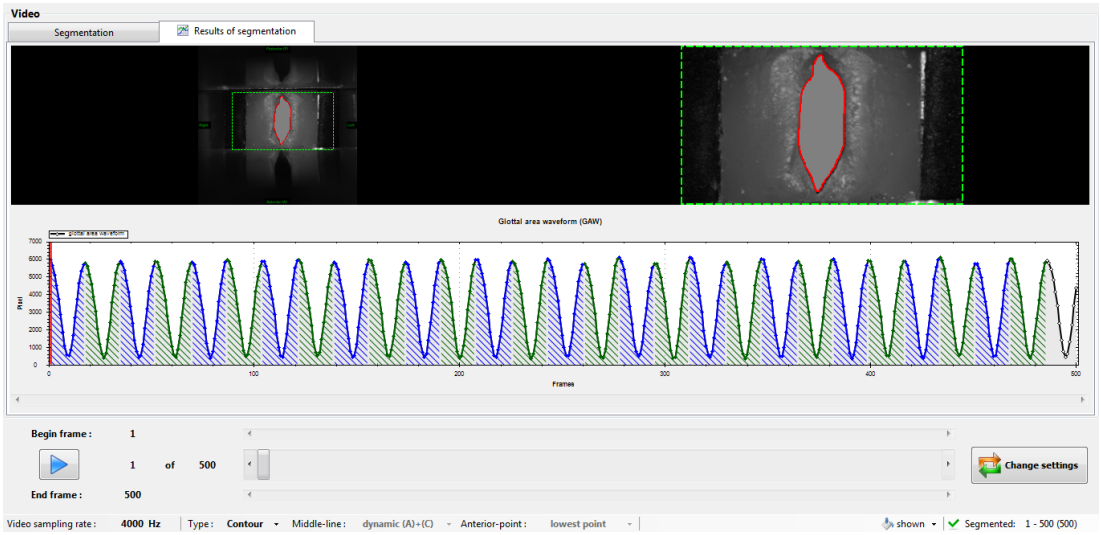
A Figure 69 - Model C114 in 18mm channel configuration in M1 during the setting of the maximum glottal gap with seatpoints with GAT.



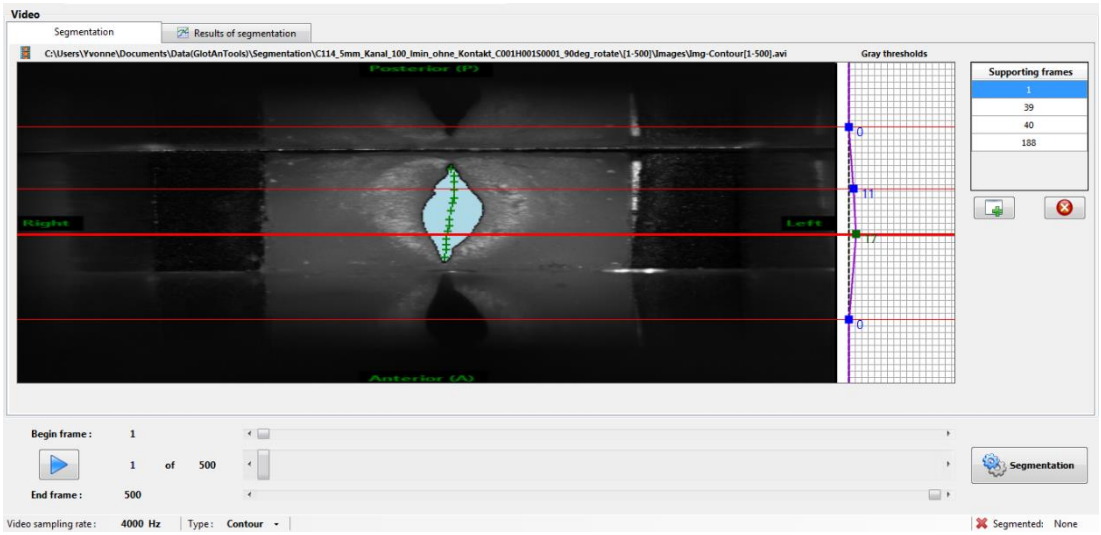
A Figure 70 - Segmentation of the model C114 in 28mm channel configuration M2 with GAT.



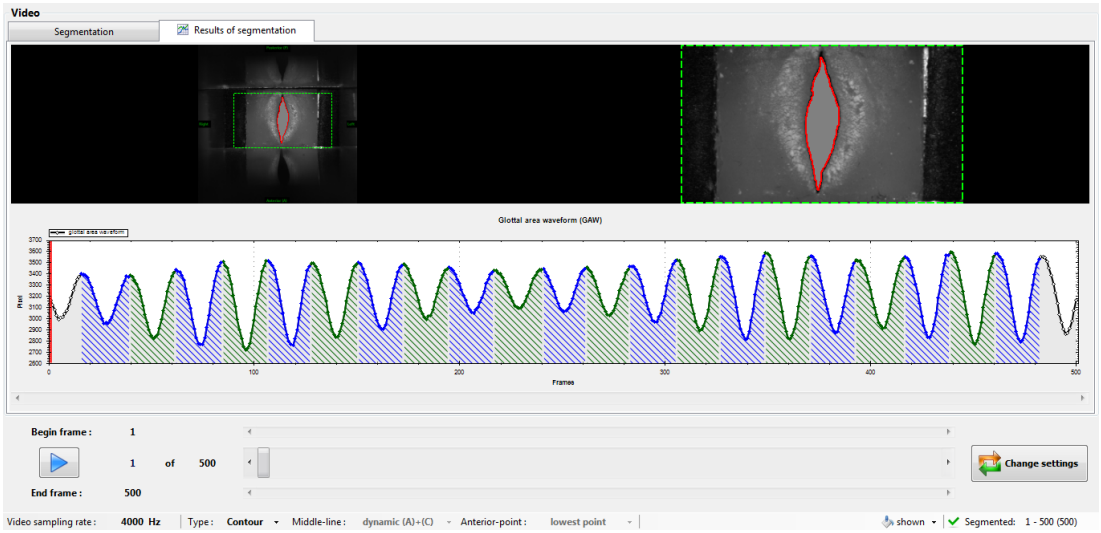
A Figure 71 - Model 114 in 28mm channel configuration in M2 during the setting of the maximum glottal gap with seatpoints with GAT.



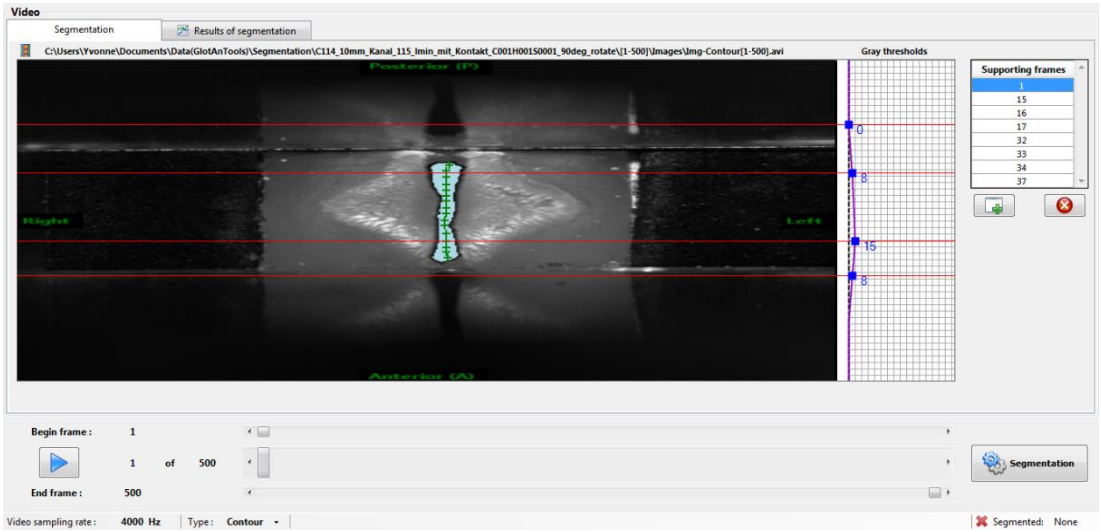
A Figure 72 - Segmentation of the model C114 in 28mm channel configuration M2 with GAT.



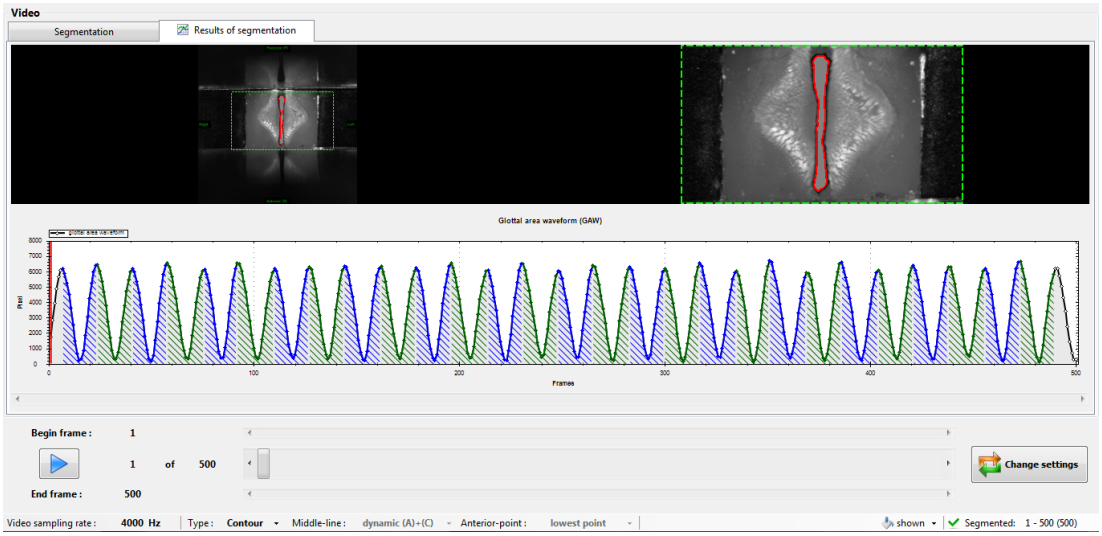
A Figure 73 - Model C114 in 28mm channel configuration in M1 during the setting of the maximum glottal gap with seatpoints with GAT.



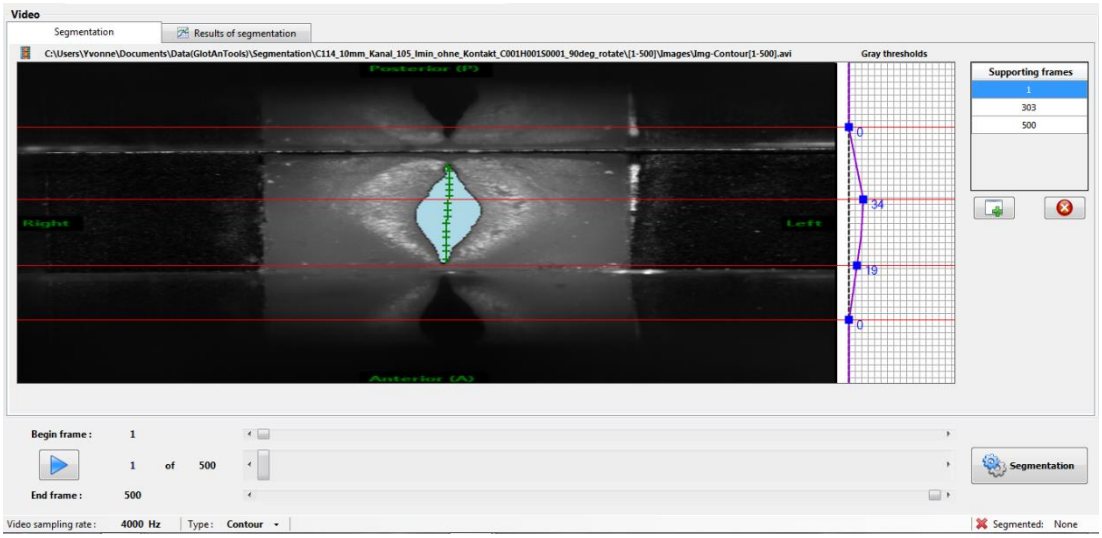
A Figure 74 - Segmentation of the model C114 in 28mm channel configuration M1 with GAT.



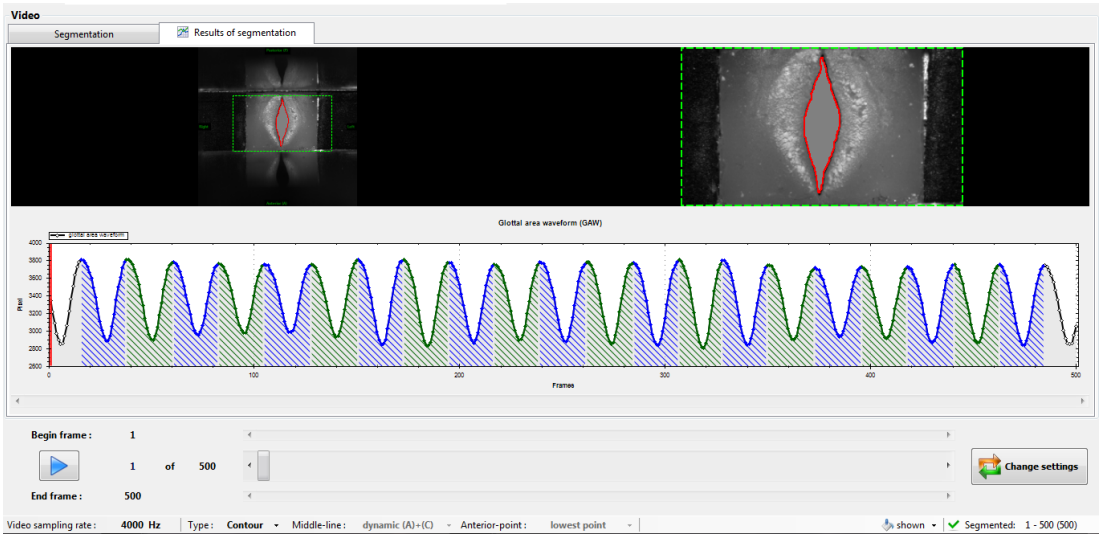
A Figure 75 - Model C114 in 38mm channel configuration in M2 during the setting of the maximum glottal gap with seatpoints with GAT.



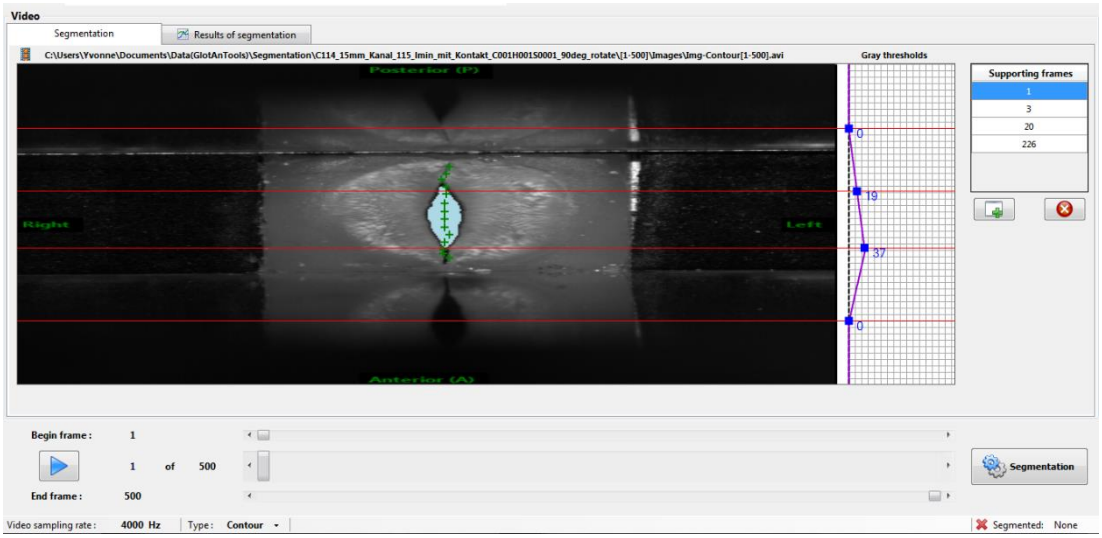
A Figure 76 - Segmentation of the model C114 in 38mm channel configuration M2 with GAT.



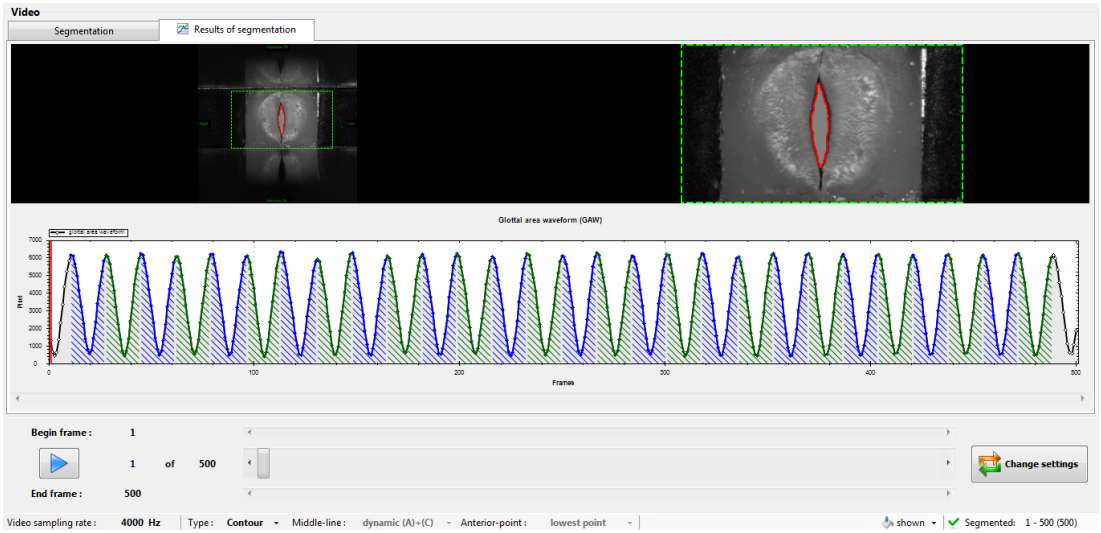
A Figure 77 - Model C114 in 38mm channel configuration in M1 during the setting of the maximum glottal gap with seatpoints with GAT.



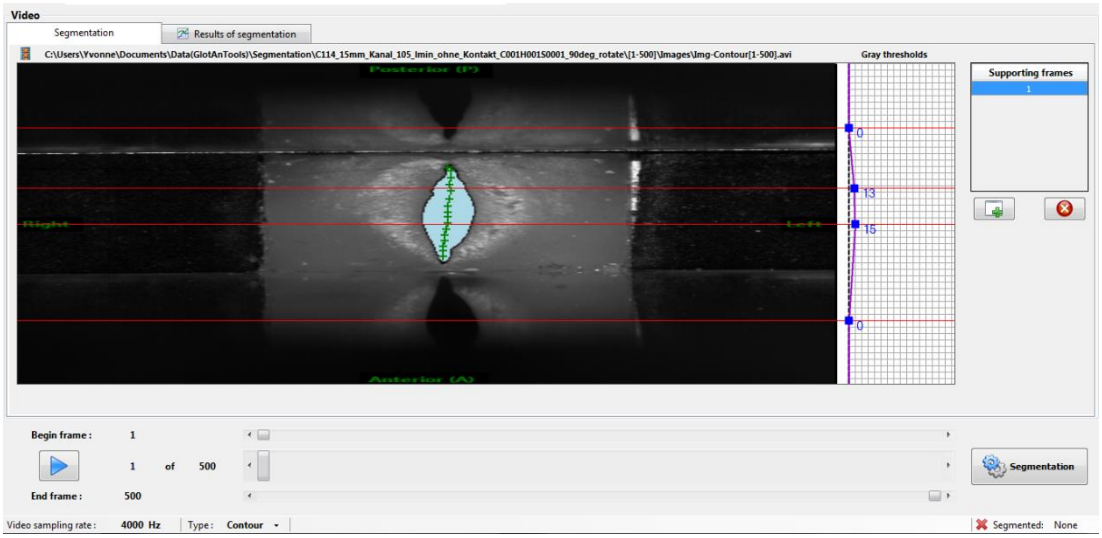
A Figure 78 - Segmentation of the model C114 in 38mm channel configuration M1 with GAT.



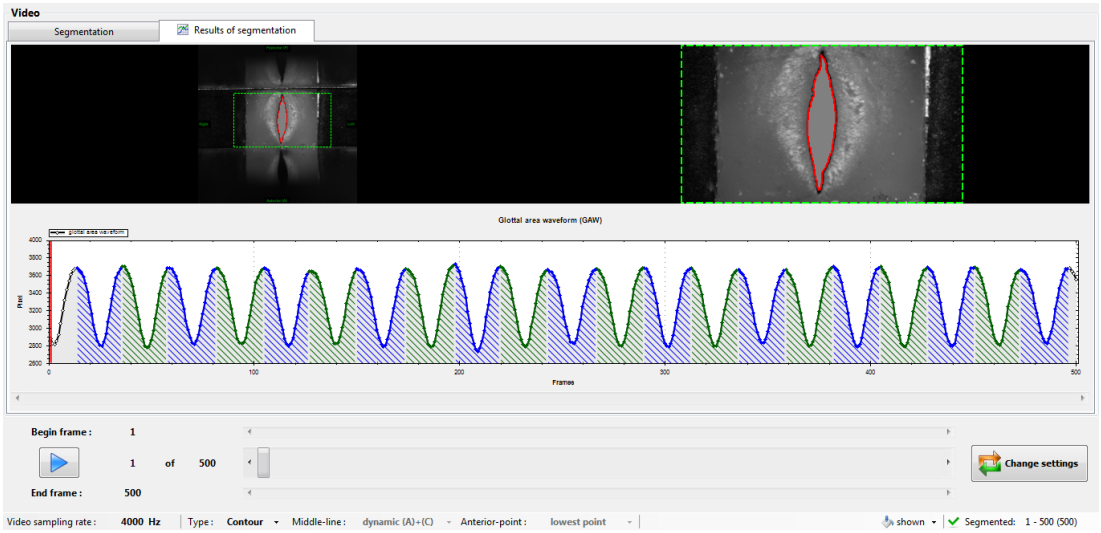
A Figure 79 - Model C114 in 48mm channel configuration in M2 during the setting of the maximum glottal gap with seatpoints with GAT.



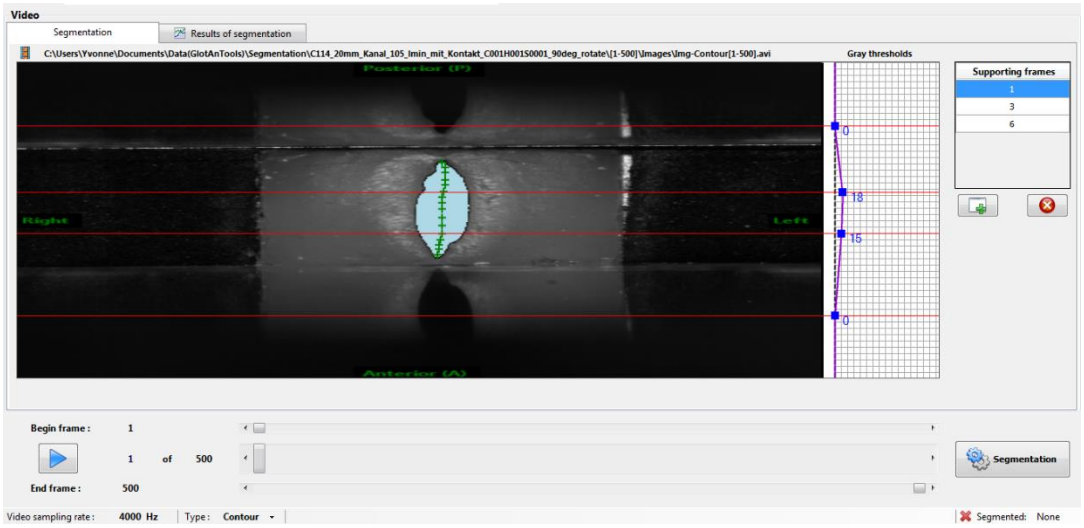
A Figure 80 - Segmentation of the model C114 in 48mm channel configuration M2 with GAT.



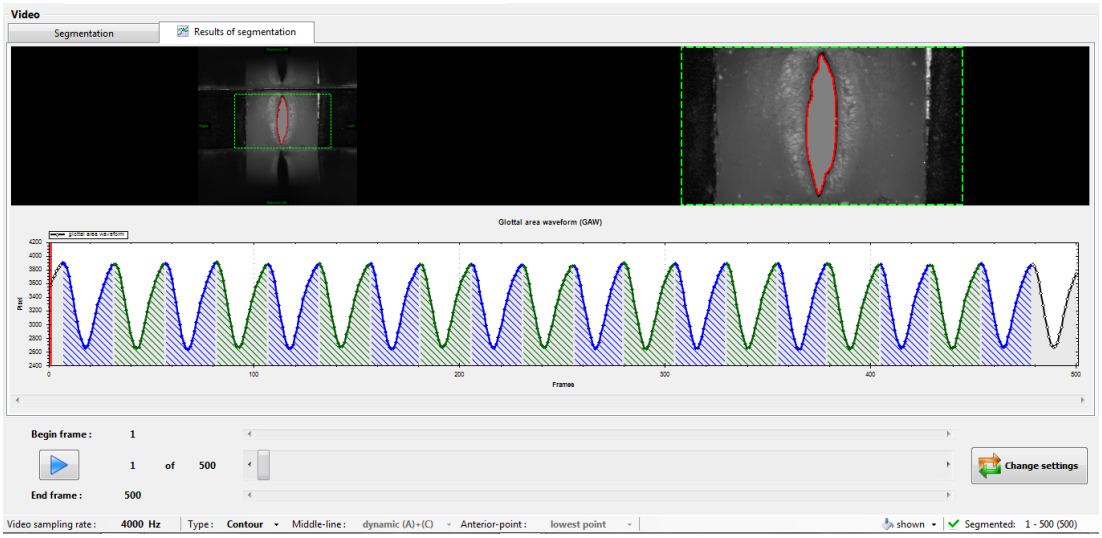
A Figure 81 - Model C114 in 48mm channel configuration in M1 during the setting of the maximum glottal gap with seatpoints with GAT.



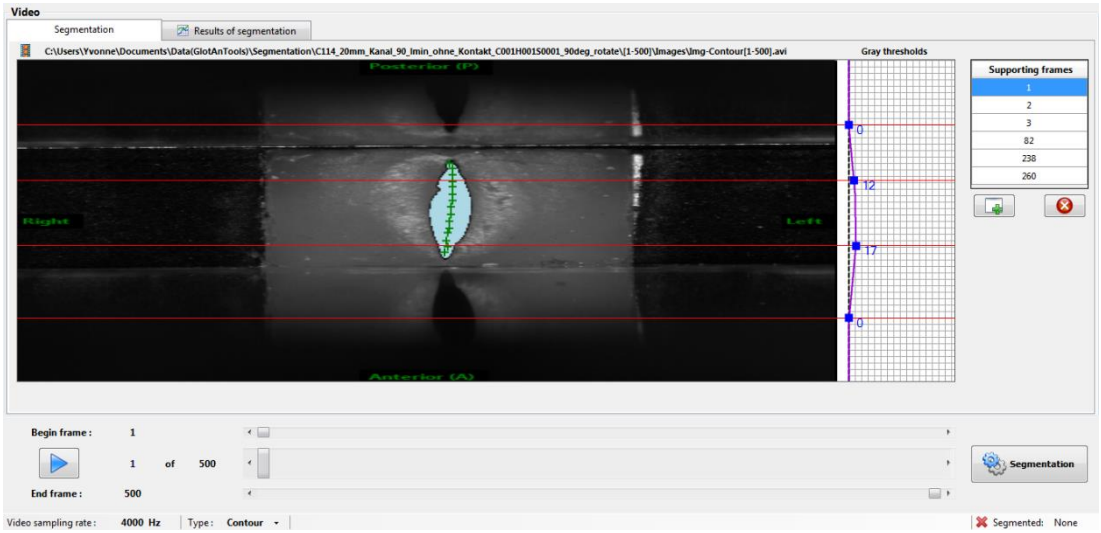
A Figure 82 - Segmentation of the model C114 in 48mm channel configuration M1 with GAT.



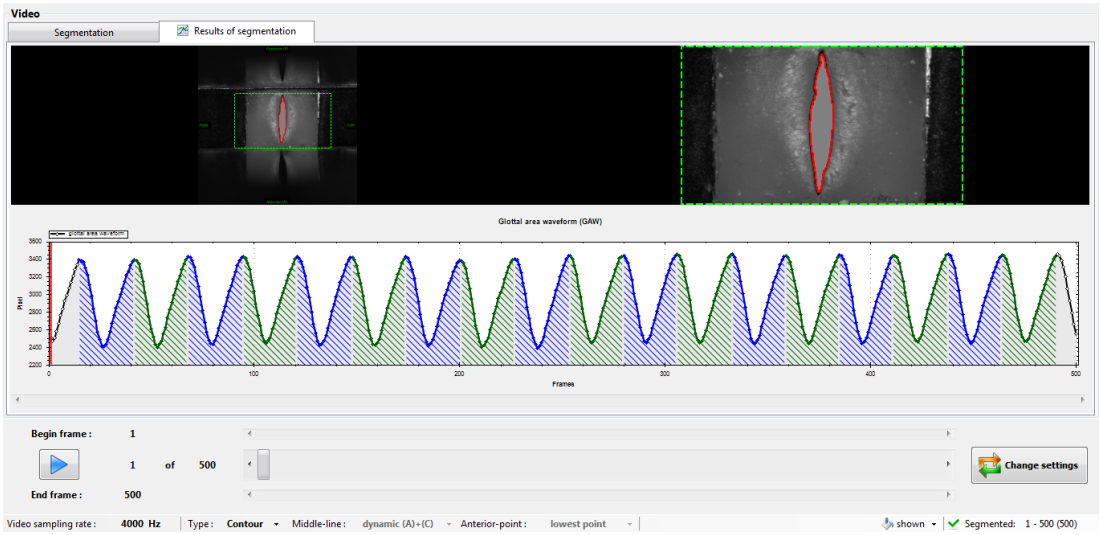
A Figure 83 - Model C114 in 58mm channel configuration in M2 during the setting of the maximum glottal gap with seatpoints with GAT.



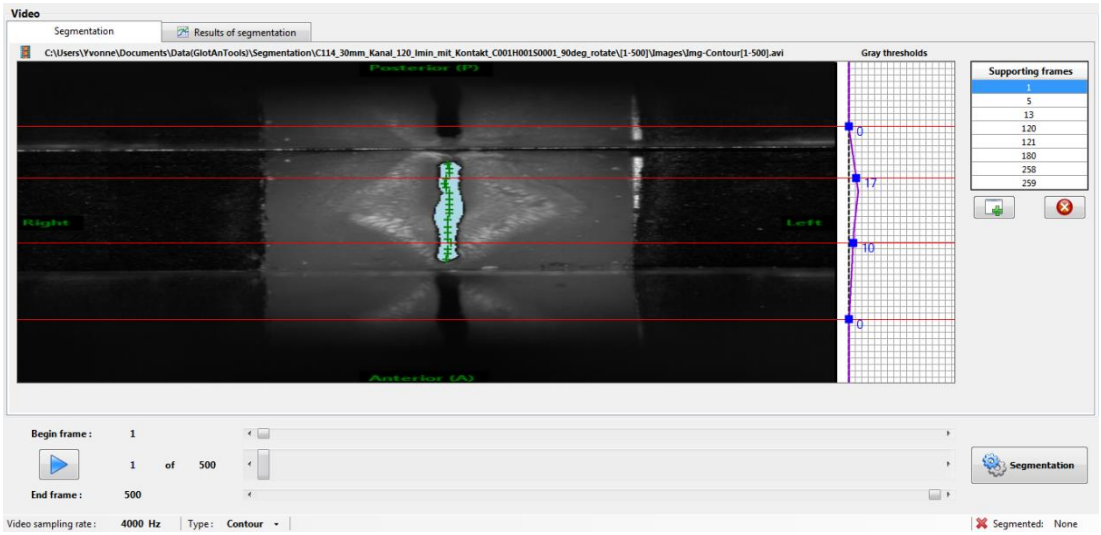
A Figure 84 - Segmentation of the model C114 in 58mm channel configuration M2 with GAT.



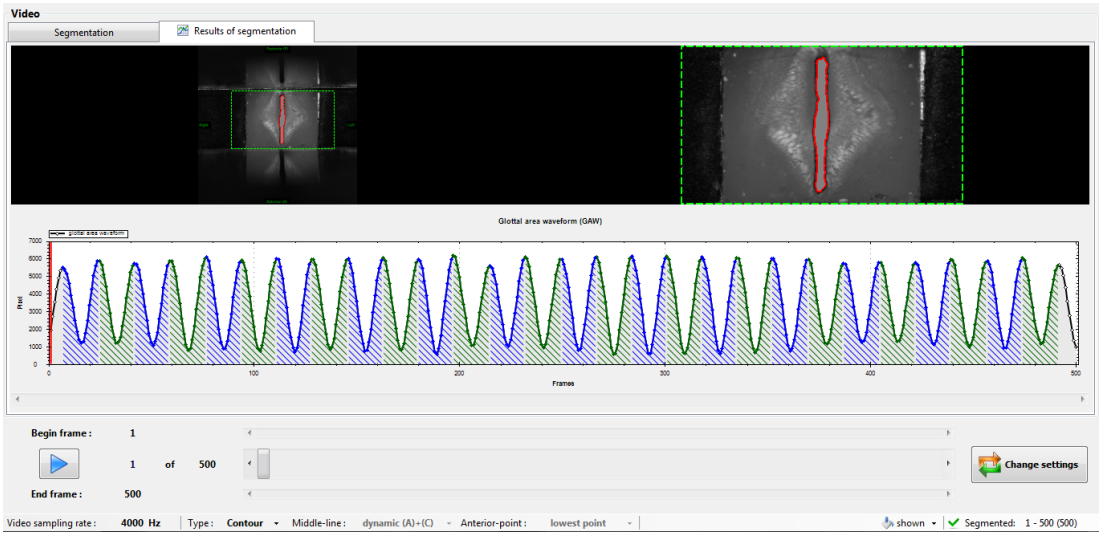
A Figure 85 - Model C114 in 58mm channel configuration in M1 during the setting of the maximum glottal gap with seatpoints with GAT.



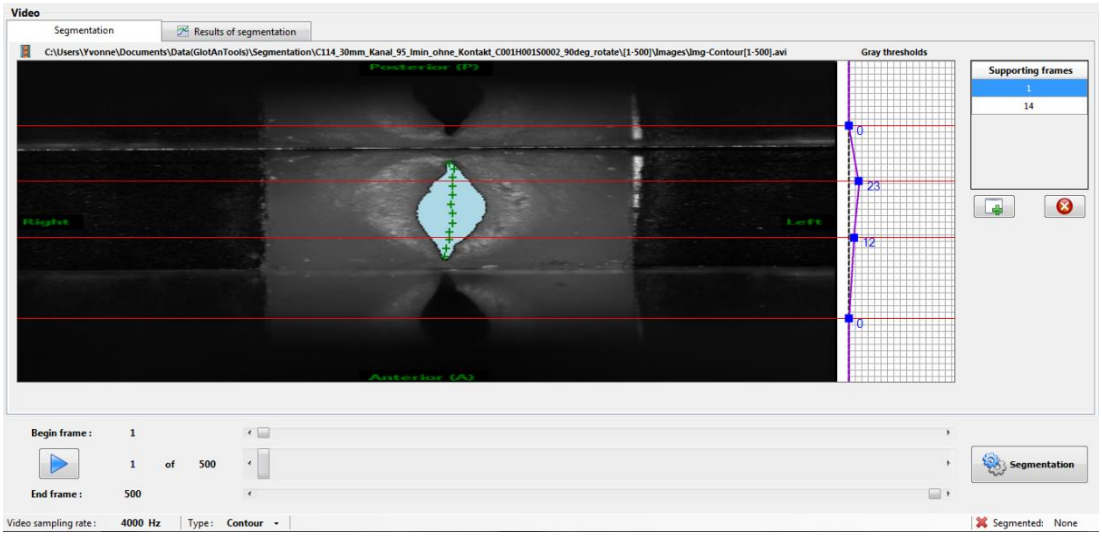
A Figure 86 - Segmentation of the model C114 in 58mm channel configuration M1 with GAT.



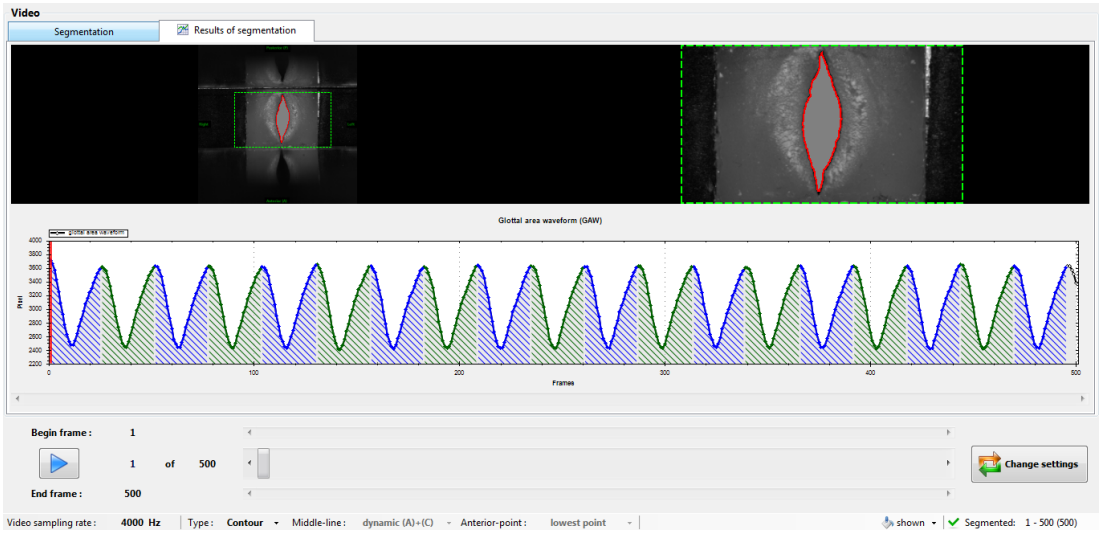
A Figure 87 - Model C114 in 78mm channel configuration in M2 during the setting of the maximum glottal gap with seatpoints with GAT.



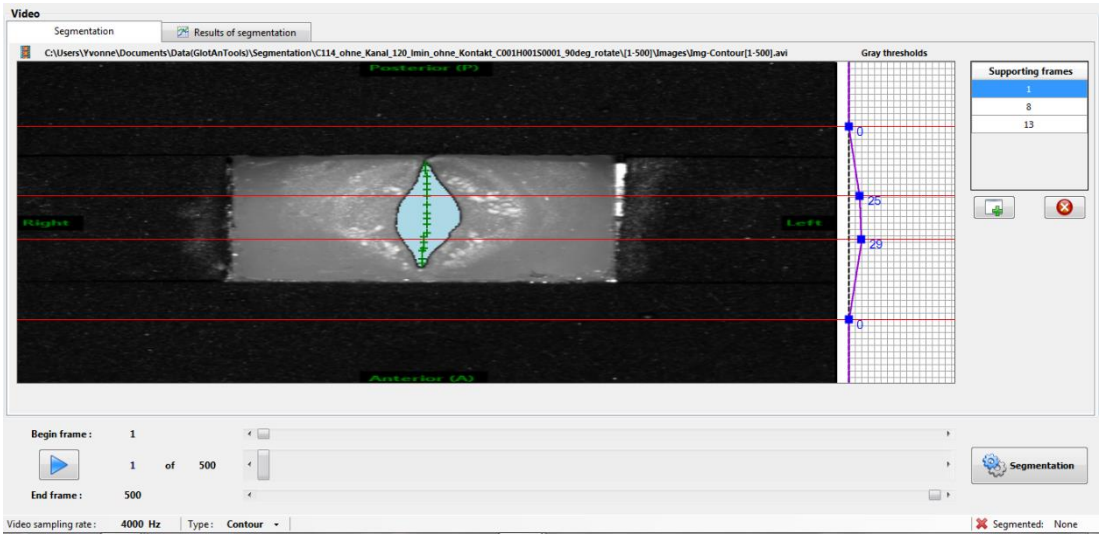
A Figure 88 - Segmentation of the model C114 in 78mm channel configuration M2 with GAT.



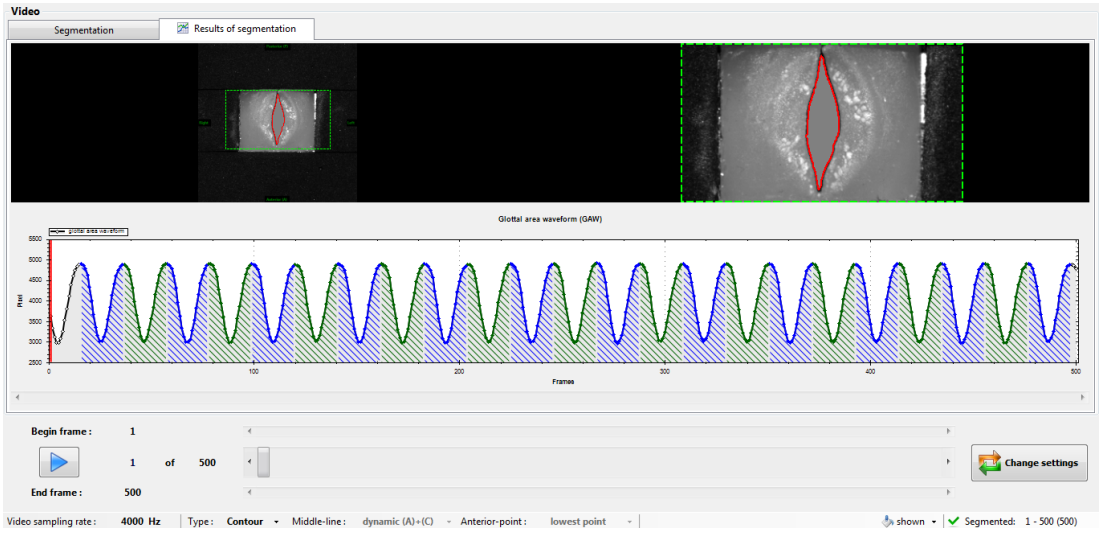
A Figure 89 - Model C114 in 78mm channel configuration in M1 during the setting of the maximum glottal gap with seatpoints with GAT.



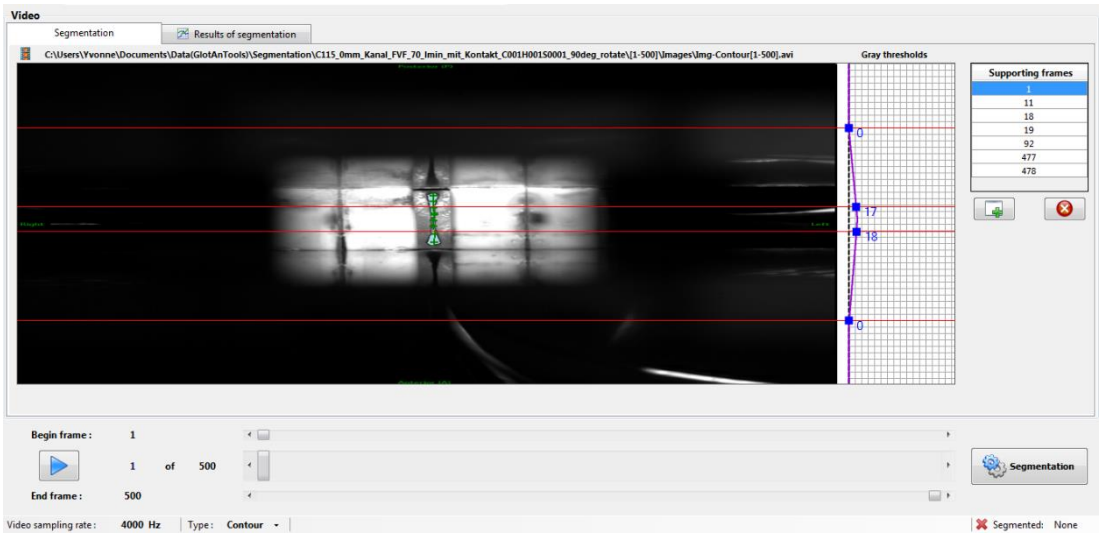
A Figure 90 - Segmentation of the model C114 in 78mm channel configuration M1 with GAT.



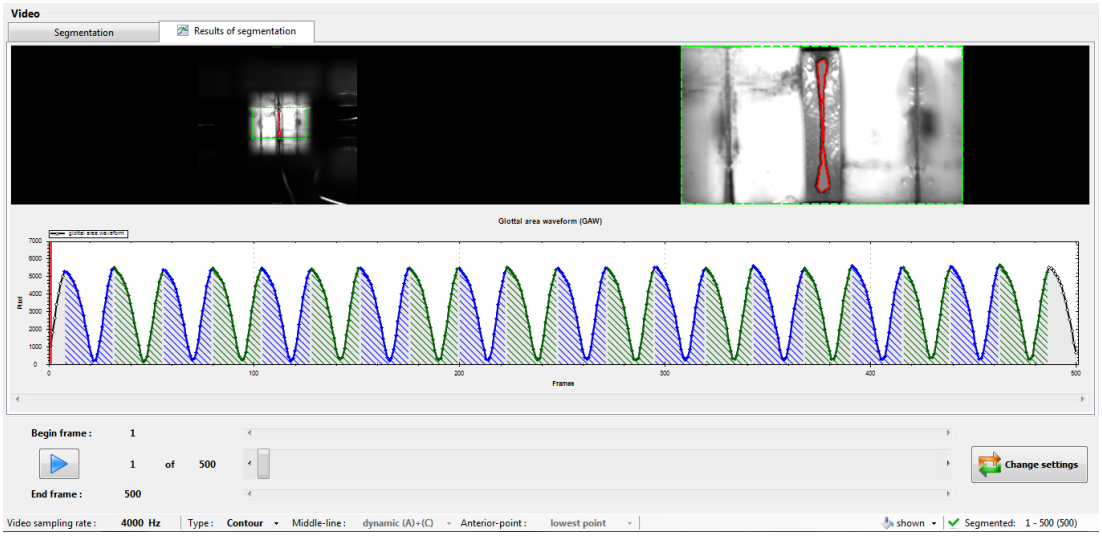
A Figure 91 - Model C114 in no channel channel configuration in M1 during the setting of the maximum glottal gap with seatpoints with GAT.



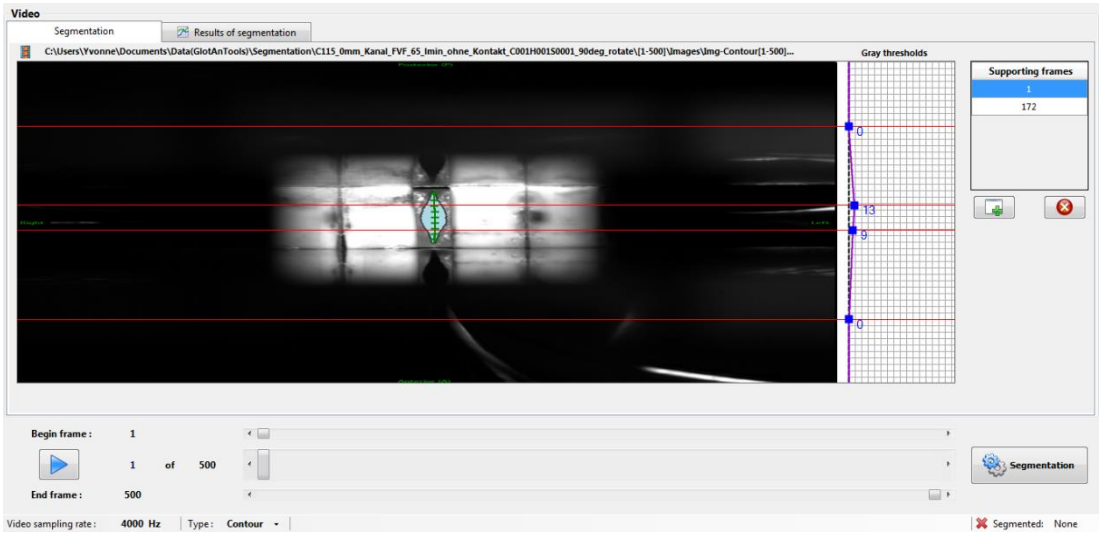
A Figure 92 - Segmentation of the model C114 in no channel channel configuration M1 with GAT.



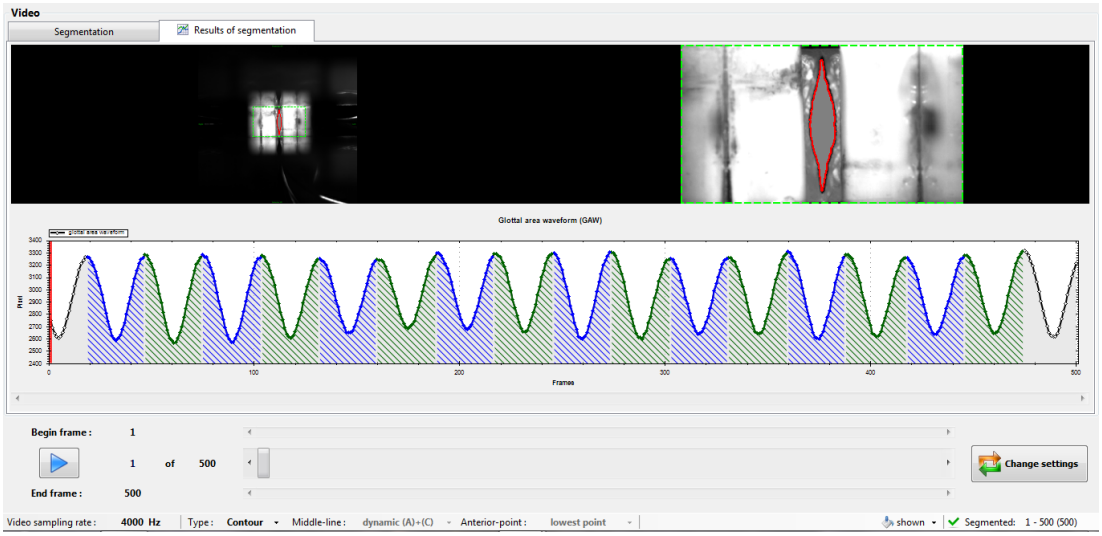
A Figure 93 - Model C115 in 18mm channel with FVF configuration in M2 during the setting of the maximum glottal gap with seatpoints with GAT.



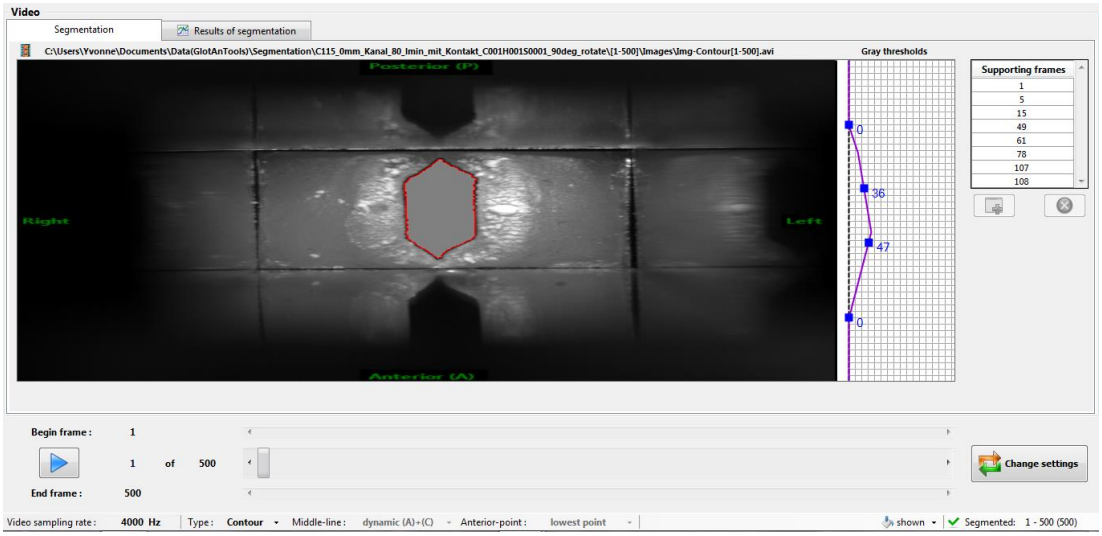
A Figure 94 - Segmentation of the model C115 in 18mm channel with FVF configuration M2 with GAT.



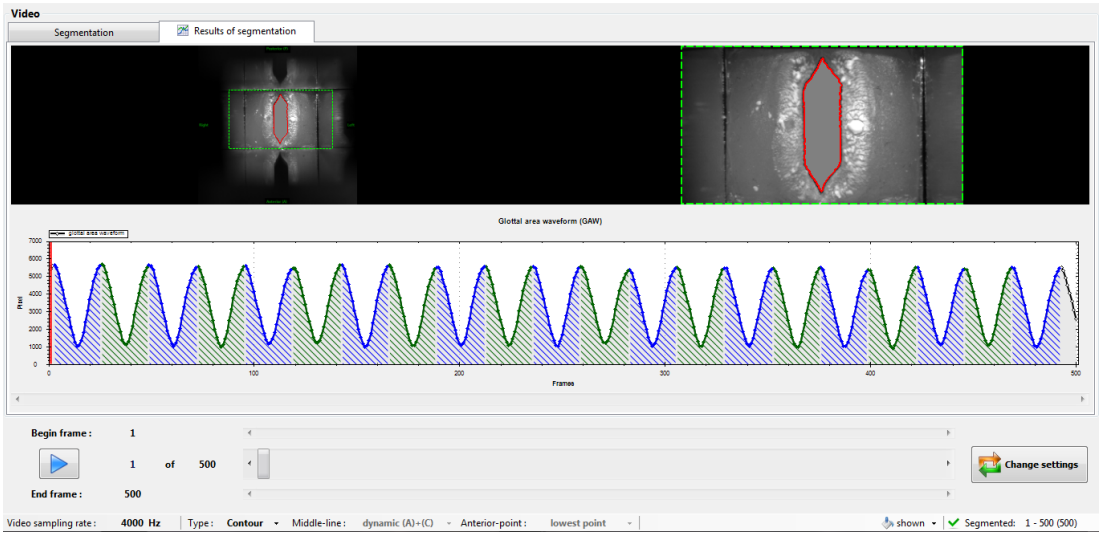
A Figure 95 - Model C115 in 18mm channel with FVF configuration in M1 during the setting of the maximum glottal gap with seatpoints with GAT.



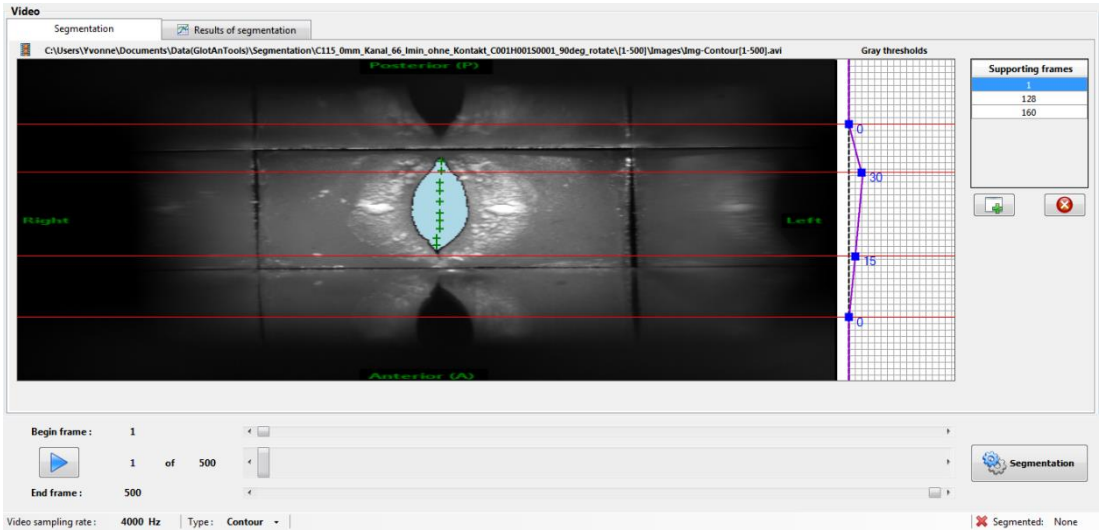
A Figure 96 - Segmentation of the model C115 in 18mm channel with FVF configuration M1 with GAT.



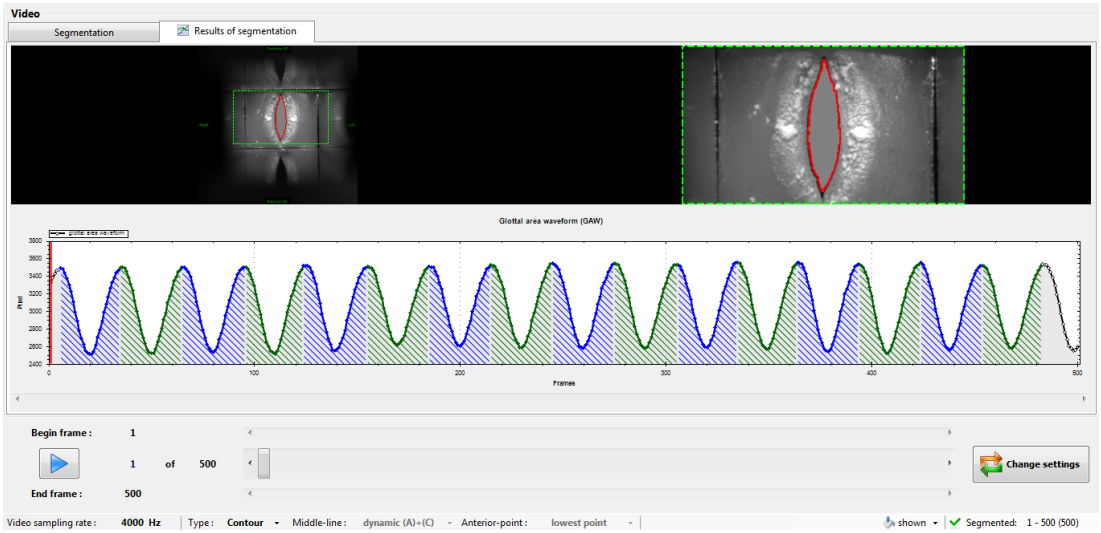
A Figure 97 - Model C115 in 18mm channel configuration in M2 during the setting of the maximum glottal gap with seatpoints with GAT.



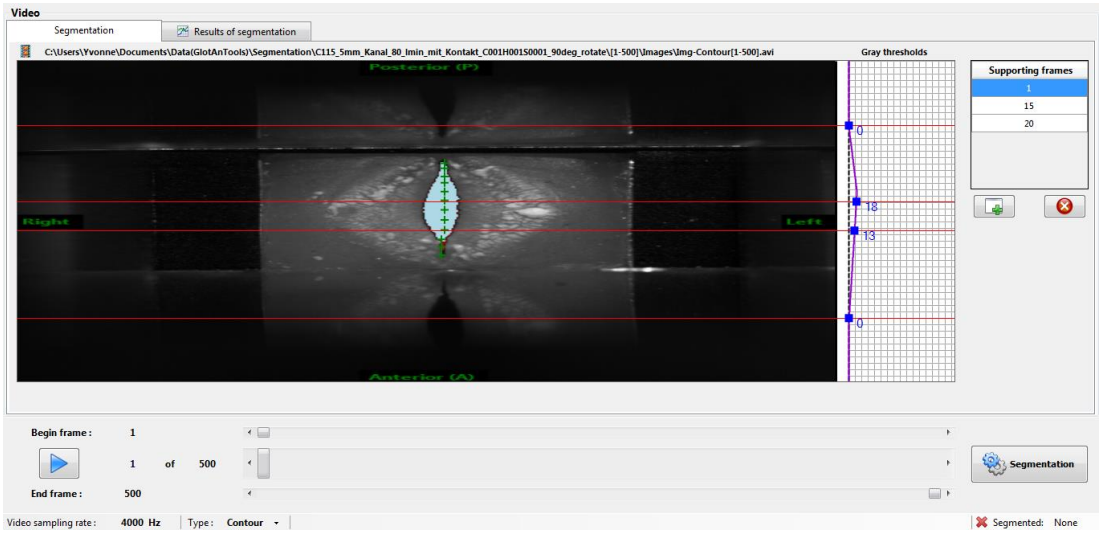
A Figure 98 - Segmentation of the model C115 in 18mm channel configuration M2 with GAT.



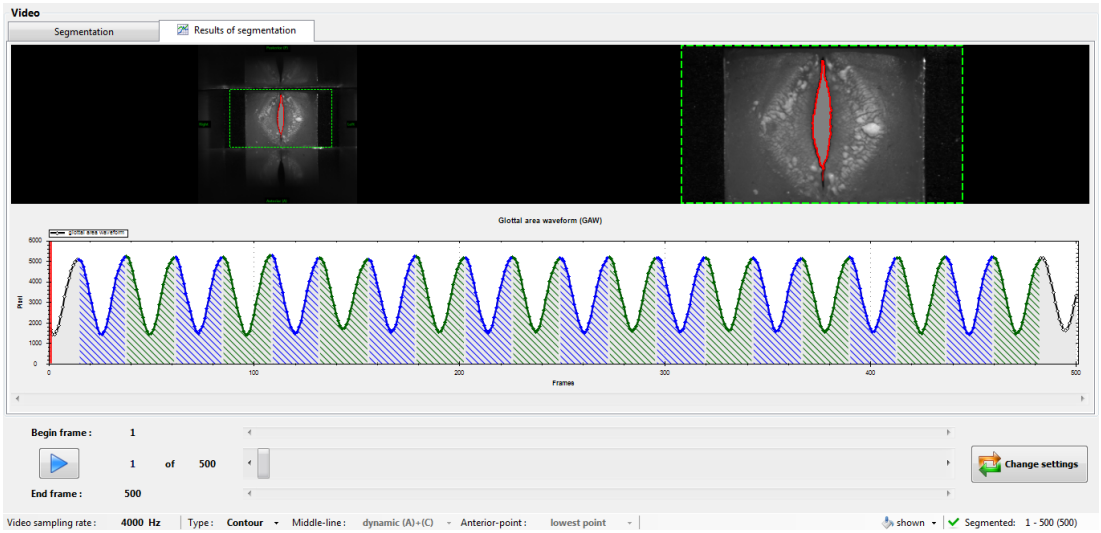
A Figure 99 - Model C115 in 18mm channel configuration in M1 during the setting of the maximum glottal gap with seatpoints with GAT.



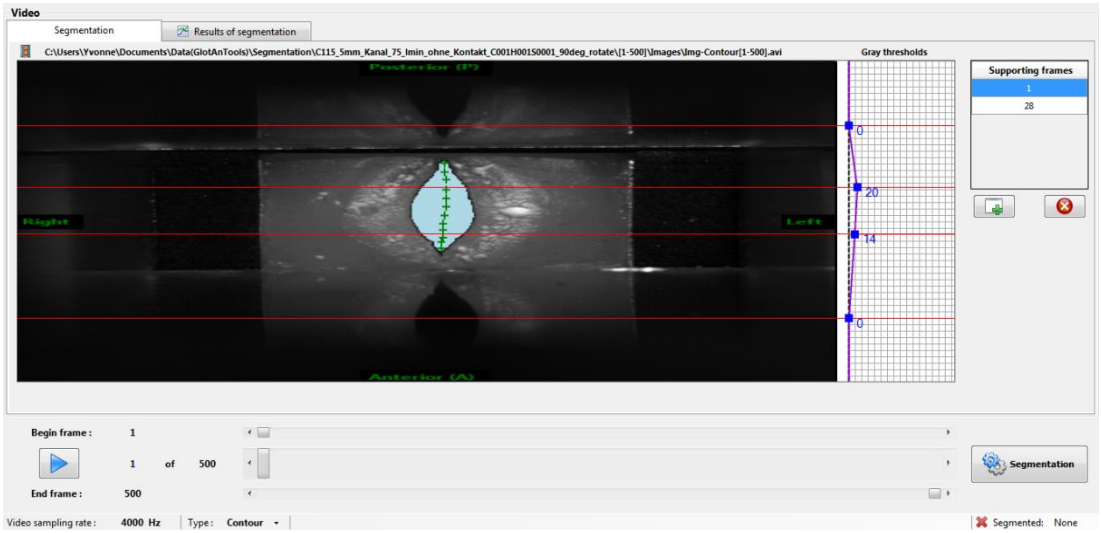
A Figure 100 - Segmentation of the model C115 in 18mm channel configuration M1 with GAT.



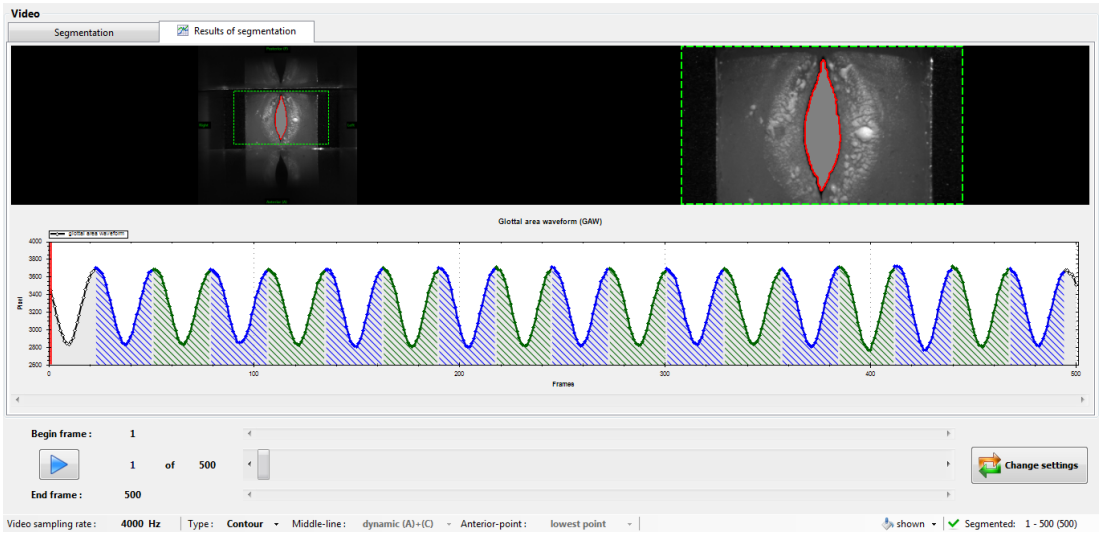
A Figure 101 - Model C115 in 28mm channel configuration in M2 during the setting of the maximum glottal gap with seatpoints with GAT.



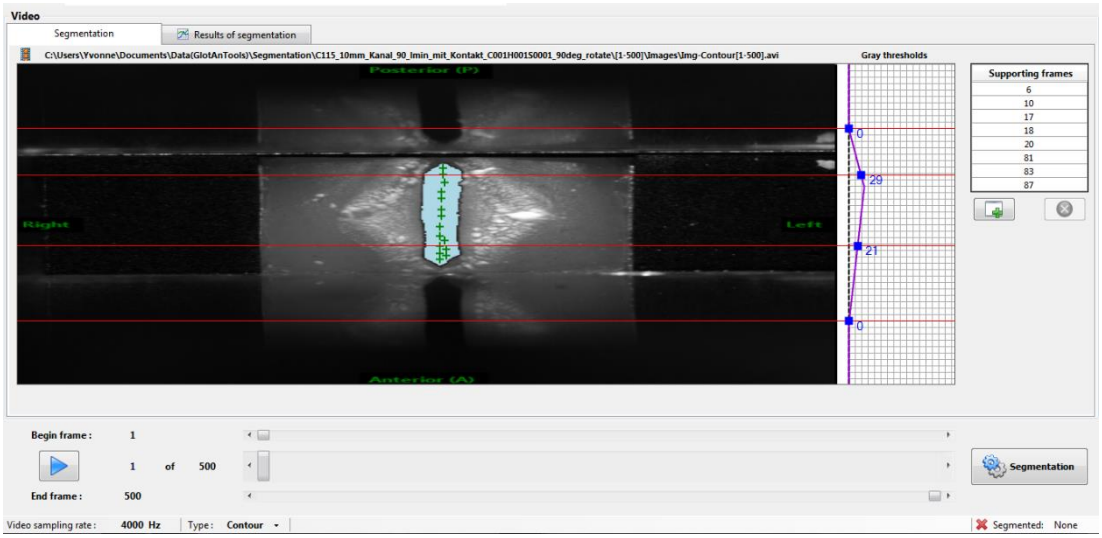
A Figure 102 - Segmentation of the model C115 in 28mm channel configuration M2 with GAT.



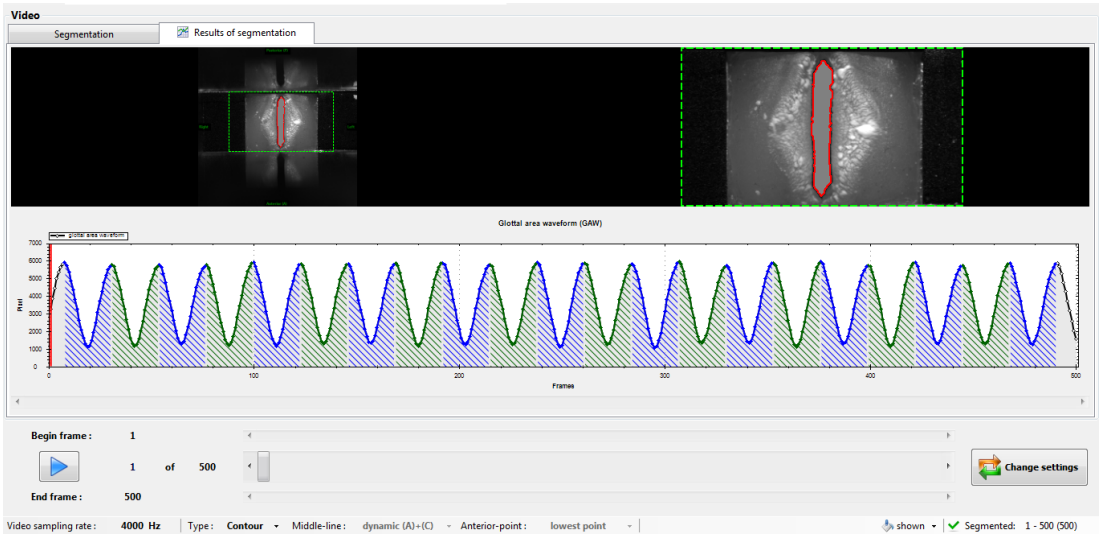
A Figure 103 - Model C115 in 28mm channel configuration in M1 during the setting of the maximum glottal gap with seatpoints with GAT.



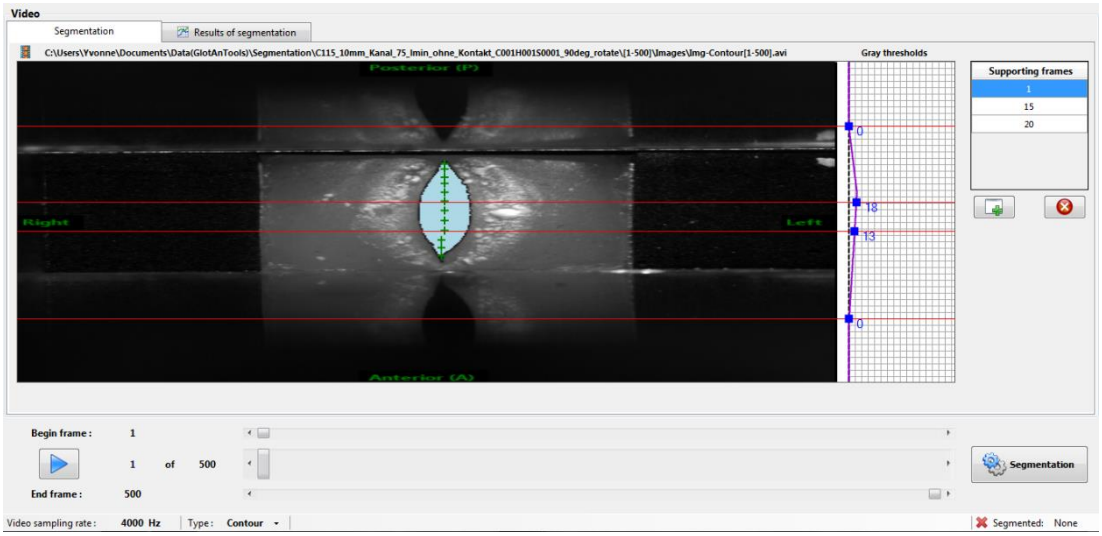
A Figure 104 - Segmentation of the model C115 in 28mm channel configuration M1 with GAT.



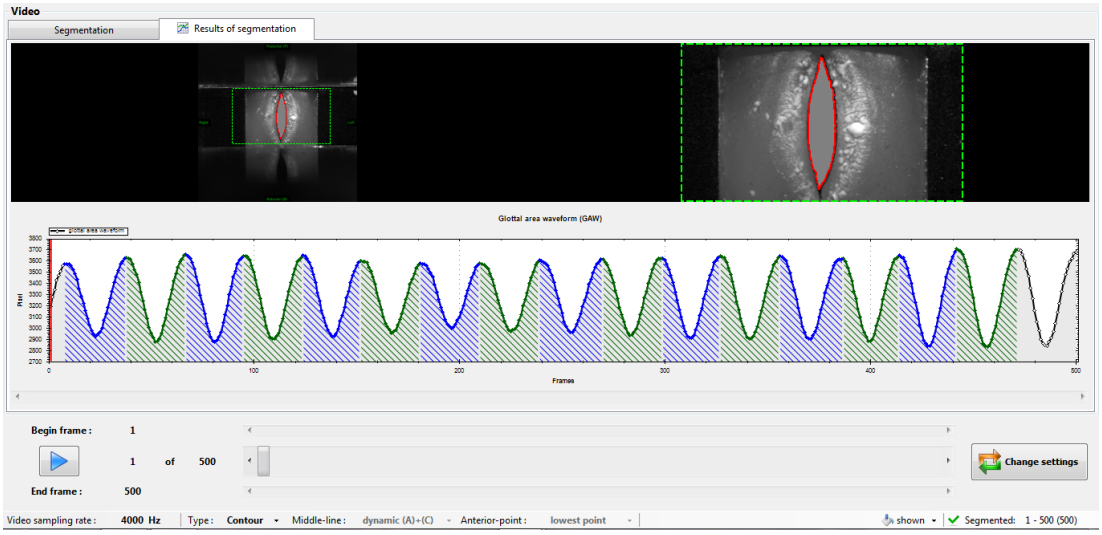
A Figure 105 - Model C115 in 38mm channel configuration in M2 during the setting of the maximum glottal gap with seatpoints with GAT.



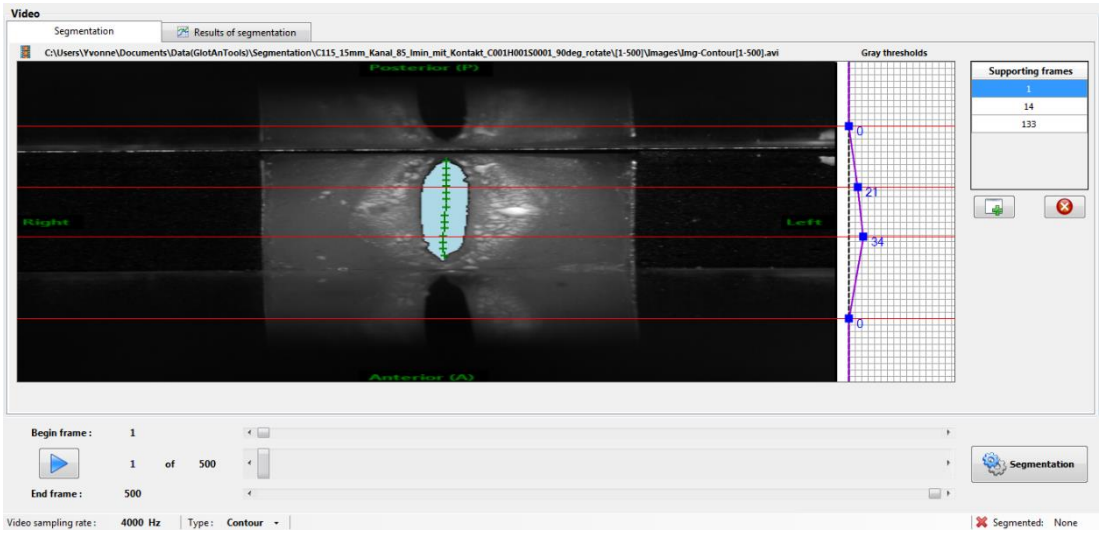
A Figure 106 - Segmentation of the model C115 in 38mm channel configuration M2 with GAT.



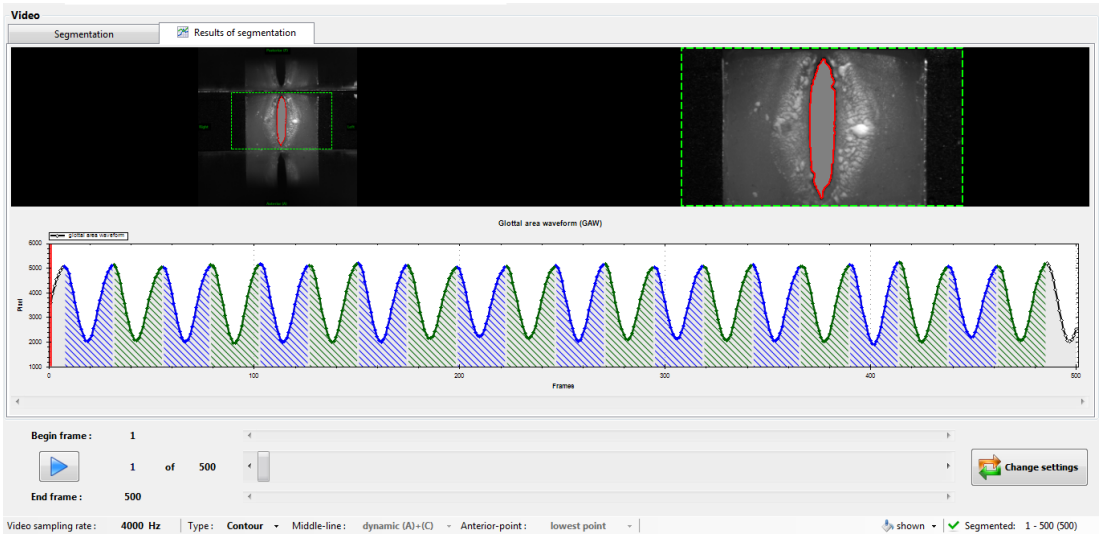
A Figure 107 - Model C115 in 38mm channel configuration in M1 during the setting of the maximum glottal gap with seatpoints with GAT.



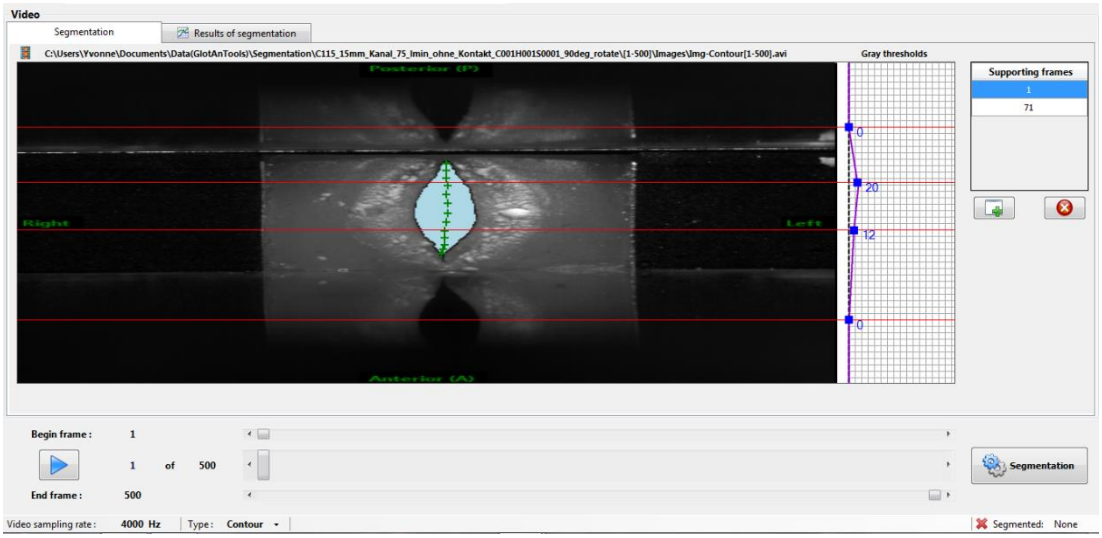
A Figure 108 - Segmentation of the model C115 in 38mm channel configuration M1 with GAT.



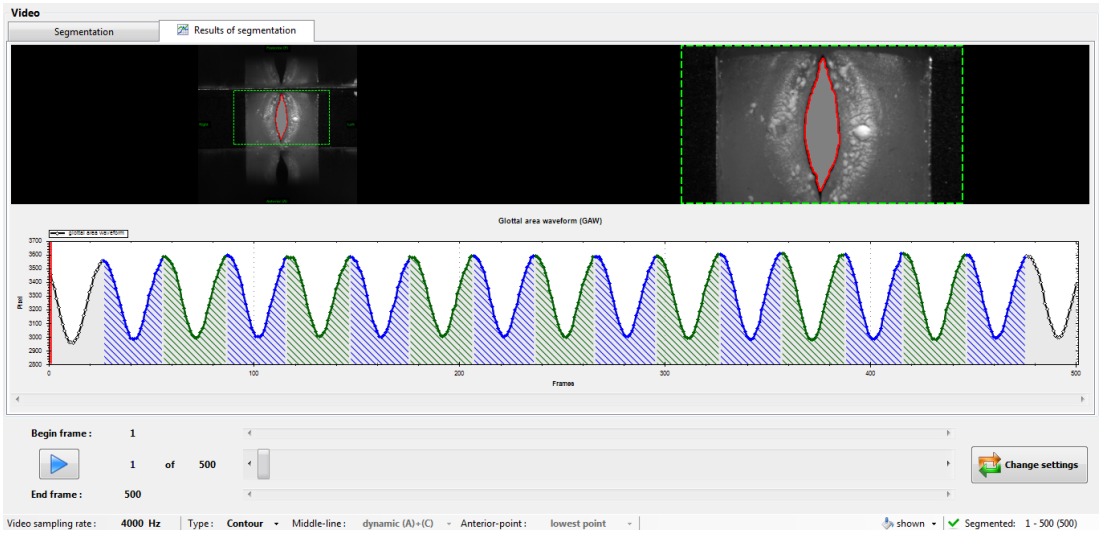
A Figure 109 - Model C115 in 48mm channel configuration in M2 during the setting of the maximum glottal gap with seatpoints with GAT.



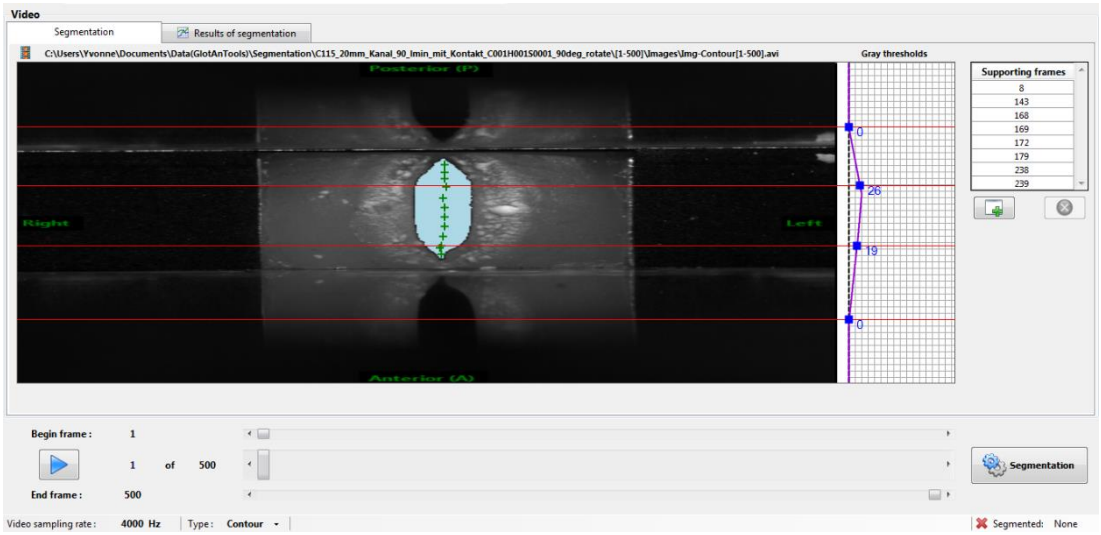
A Figure 110 - Segmentation of the model C115 in 48mm channel configuration M2 with GAT.



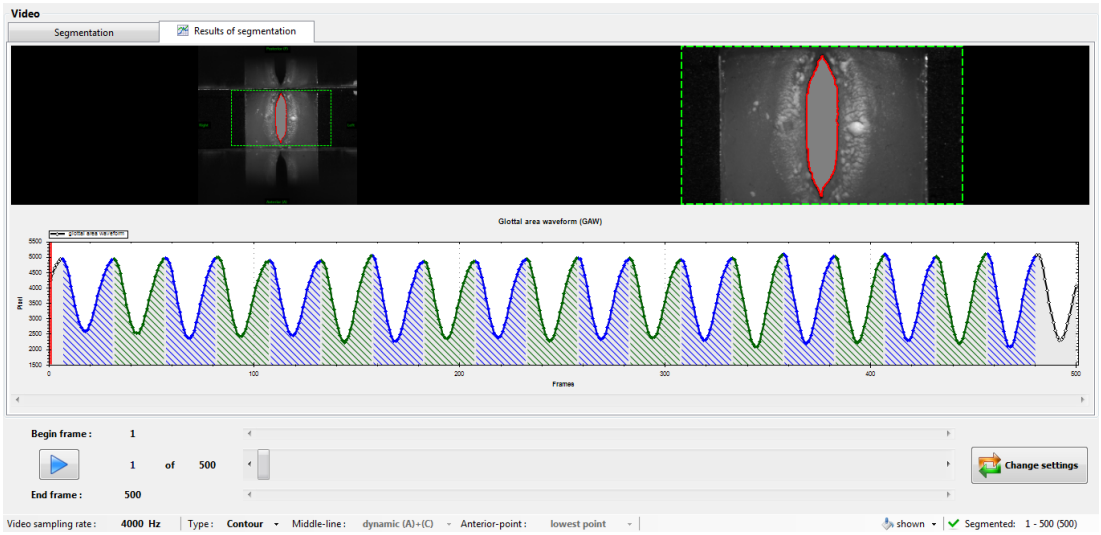
A Figure 111 - Model C115 in 48mm channel configuration in M1 during the setting of the maximum glottal gap with seatpoints with GAT.



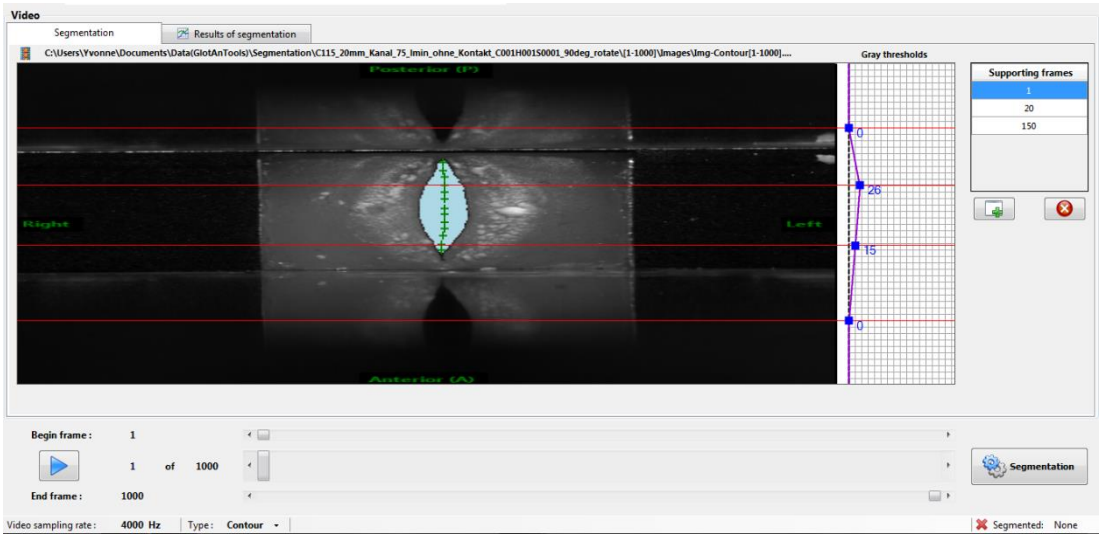
A Figure 112 - Segmentation of the model C115 in 48mm channel configuration M1 with GAT.



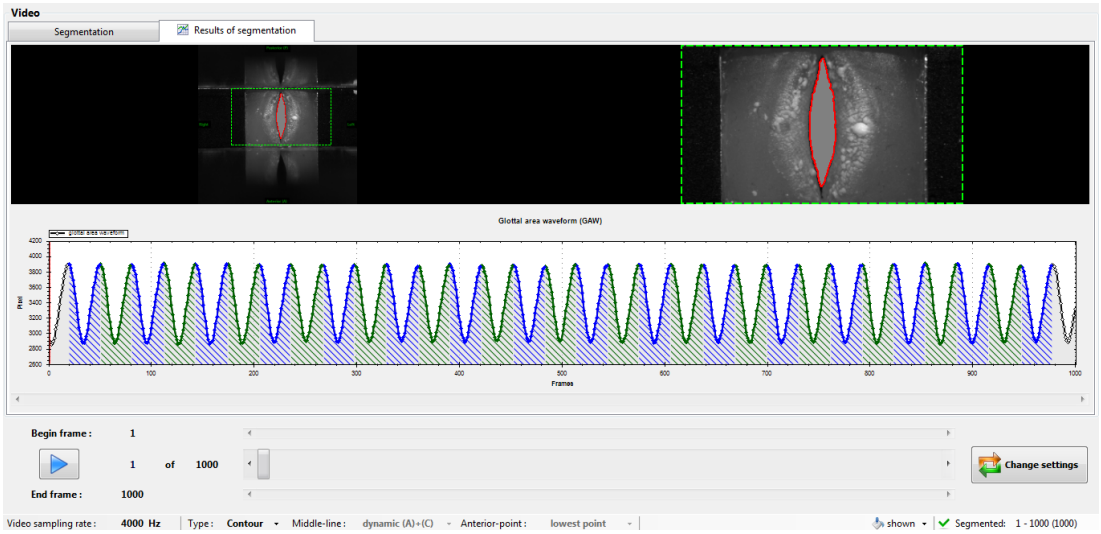
A Figure 113 - Model C115 in 58mm channel configuration in M2 during the setting of the maximum glottal gap with seatpoints with GAT.



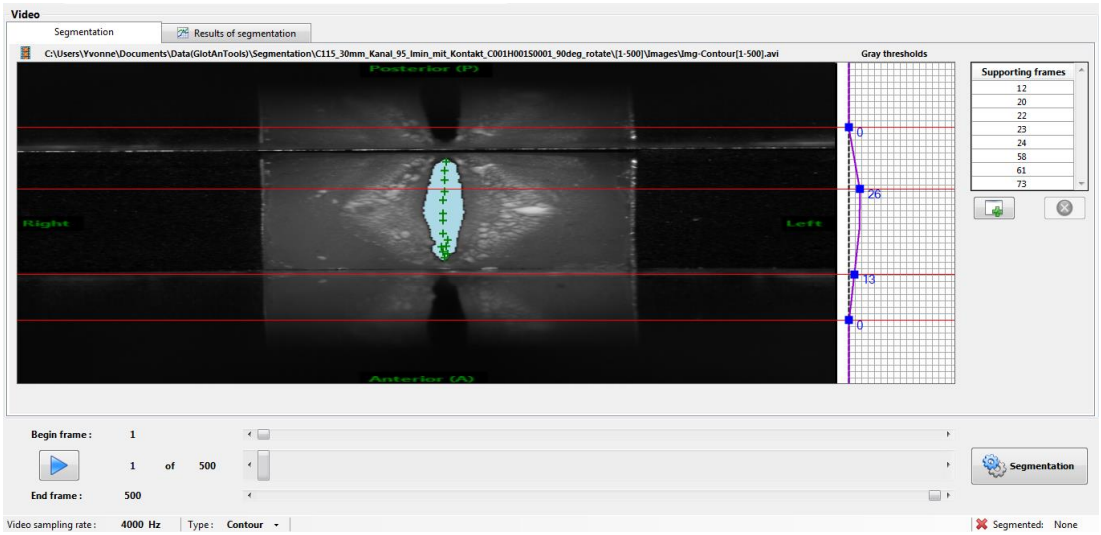
A Figure 114 - Segmentation of the model C115 in 58mm channel configuration M2 with GAT.



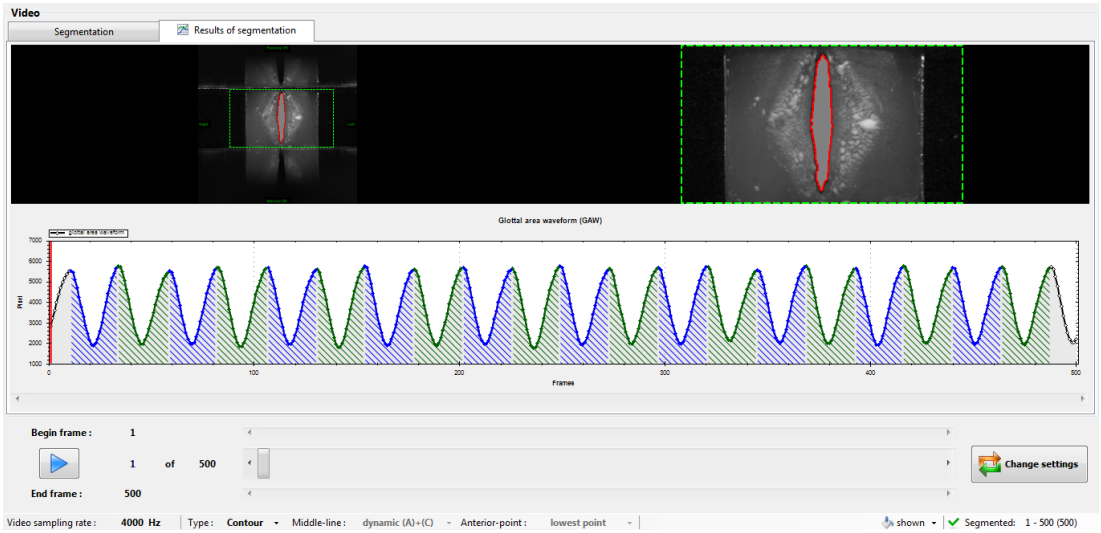
A Figure 115 - Model C115 in 58mm channel configuration in M1 during the setting of the maximum glottal gap with seatpoints with GAT.



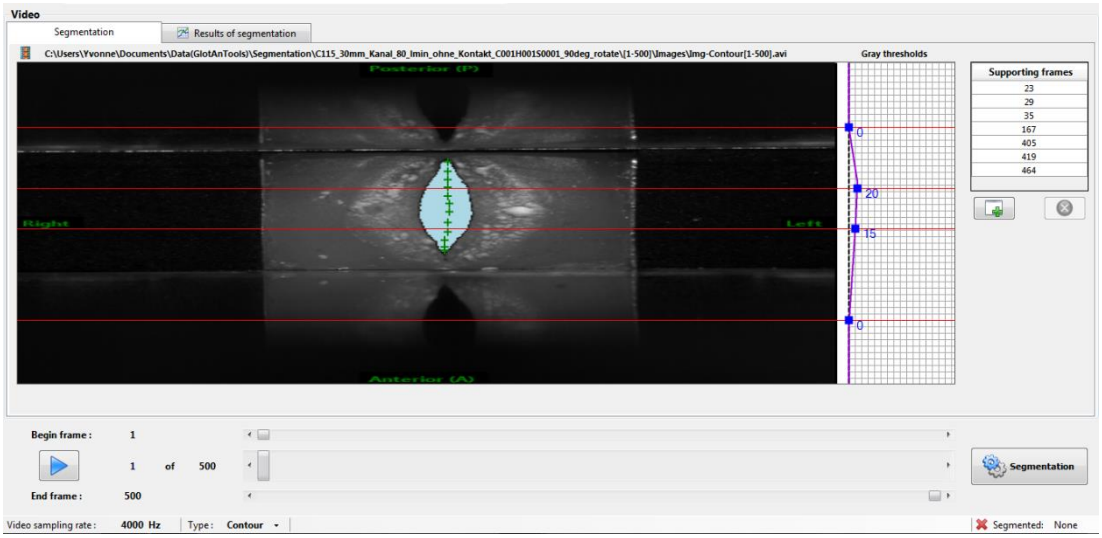
A Figure 116 - Segmentation of the model C115 in 58mm channel configuration M1 with GAT.



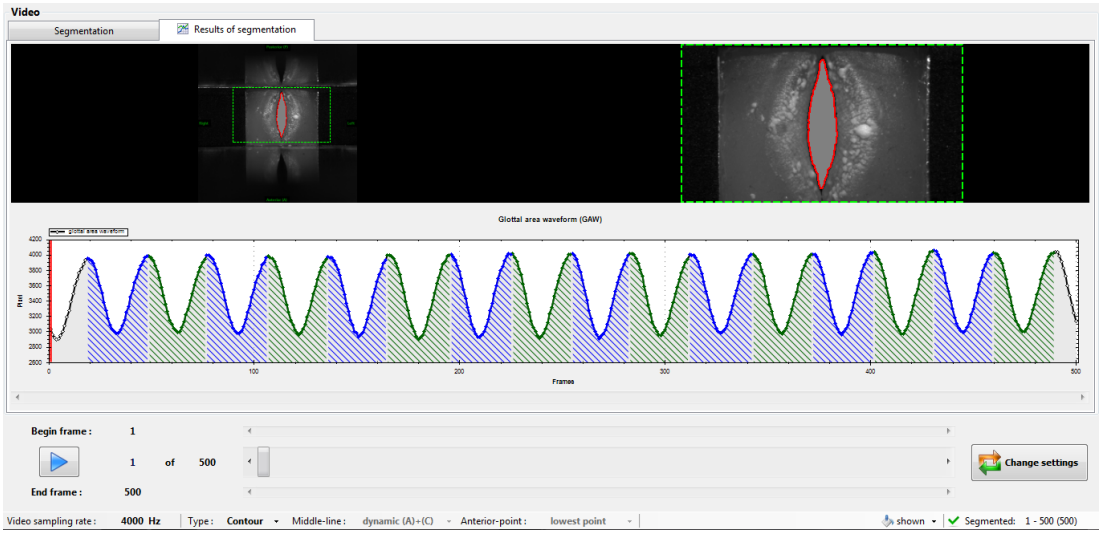
A Figure 117 - Model C115 in 78mm channel configuration in M2 during the setting of the maximum glottal gap with seatpoints with GAT.



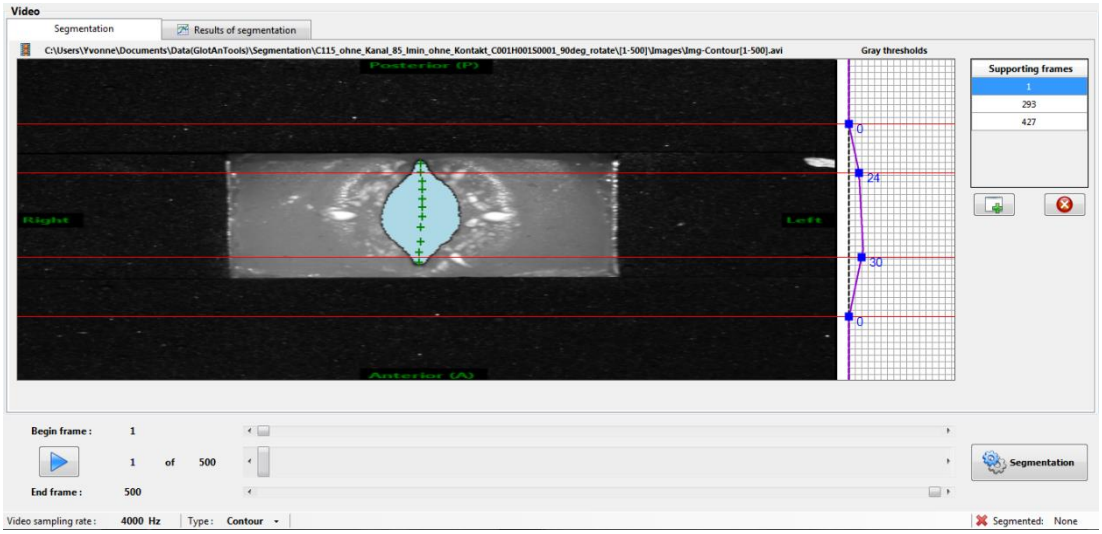
A Figure 118 - Segmentation of the model C115 in 78mm channel configuration M2 with GAT.



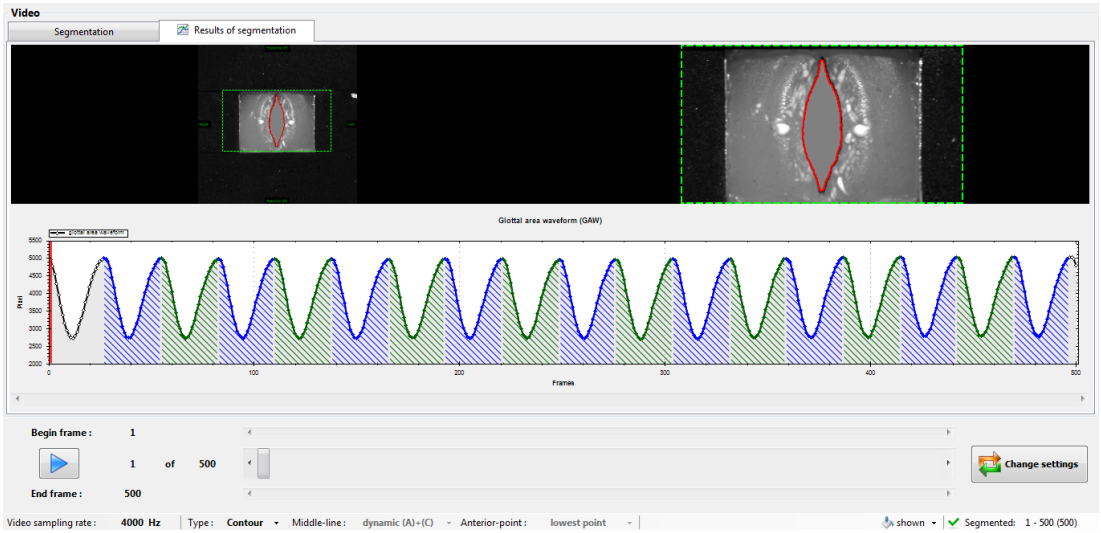
A Figure 119 - Model C115 in 78mm channel configuration in M1 during the setting of the maximum glottal gap with seatpoints with GAT.



A Figure 120 - Segmentation of the model C115 in 78mm channel configuration M1 with GAT.



A Figure 121 - Model C115 in no channel channel configuration in M1 during the setting of the maximum glottal gap with seatpoints with GAT.



A Figure 122 - Segmentation of the model C115 in no channel channel configuration M1 with GAT.