Delineating the Operational Envelope of Mobile and Conventional EDA Sensing on Key Body Locations

Panagiotis Tsiamyrtzis
Dept. of Statistics
Athens Univ. of Econ. & Bus.
pt@aueb.gr

Malcolm Dcosta
Computational Physiology Lab
Univ. Of Houston
mtdcosta2@uh.edu

Dvijesh Shastri
Computer Science
Univ. of Houston-Downtown
shastrid@uh.edu

Eswar Prasad
Computational Physiology Lab
Univ. of Houston
emanapuram@uh.edu

Ioannis Pavlidis
Computational Physiology Lab
Univ. of Houston
ipavlidis@uh.edu

ABSTRACT
Electrodermal activity (EDA) is an important affective indicator, measured conventionally on the fingers with desktop sensing instruments. Recently, a new generation of wearable, battery-powered EDA devices came into being, encouraging the migration of EDA sensing to other body locations. To investigate the implications of such sensor/location shifts in psychophysiological studies we performed a validation experiment. In this experiment we used startle stimuli to instantaneously arouse the sympathetic system of \( n = 23 \) subjects while sitting. Startle stimuli are standard but minimal stressors, and thus ideal for determining the sensor and location resolution limit. The experiment revealed that precise measurement of small EDA responses on the fingers and palm is feasible either with conventional or mobile EDA sensors. By contrast, precise measurement of small EDA responses on the sole is challenging, while on the wrist even detection of such responses is problematic for both EDA modalities. Given that affective wristbands have emerged as the dominant form of EDA sensing, researchers should beware of these limitations.

Author Keywords
Electrodermal Activity (EDA); skin conductance; affective sensors; wearable sensors; startle

ACM Classification Keywords
H.5.2 Information Interfaces and Presentation (e.g. HCI): User Interfaces; Input devices and strategies

INTRODUCTION
Sympathetic signals are fundamental indicators of emotional state in affective human studies [1]. Peripheral sensing of sympathetic signals primarily targets transient perspiratory responses, also known as electrodermal activity (EDA) [2]. Conventionally, EDA is measured with clinical instruments that are grid-powered, have limited Internet connectivity, and feature tethered probes. These probes are attached to the subject’s fingers. Such conventional sensing schemes restrict mobility, and in some cases interfere with the subject’s task (e.g., dexterous action studies [10]). The availability of more autonomous and unobtrusive forms of EDA sensing has increased with the proliferation of mobile computing [11].

There are several mobile EDA sensors in wristband or other wearable forms. These include the Q Sensor (Affectiva, Waltham, Massachusetts), the Shimmer3 GSR + Optical Pulse Development Kit (Shimmer, Dublin, Ireland), and the E4 Wristband (Empatica, Milan, Italy). There are also several conventional EDA instruments built around desktop data acquisition modules. These include the Powerlab (ADInstruments, Bella Vista, New South Wales), the MP 150 (Biopac, Goleta, California), and others.

The technology within each family of EDA instruments (conventional vs. mobile) is largely the same, but there are differences across families. Specifically, both instrument families use Ag/AgCl disc electrodes with contact areas of 1.0 cm\(^2\) for their recordings, as recommended in the literature [4]; they differ, however, in terms of power mode (AC vs. DC), packaging (large vs. small form factor), standard communication capabilities (wired vs. wireless), and wearable options. The power mode has some implications in measurement accuracy (AC is better [2]), while the wearable options in the mobile family of sensors offer flexibility but also present a challenge. The challenge stems from the fact that sympathetically induced EDA responses differ in their strength among various body locations [2, 18]. Unfortunately, the strongest responses do not necessarily correlate with the most ‘wearable’ body locations. The combination of lesser accuracy with lesser responses may bias results in mobile studies of subject affect. Such biases have not been sufficiently appreciated or studied in the literature.

M. van Dooren et al. investigated the relative strength of EDA responses on 16 body locations [18]. They aroused subjects via emotional film segments, measuring responses with a conventional EDA device. The study concluded that EDA measurements at the foot were most similar with those of the finger. In contrast, arm, back, and armpit traces differed most
from the finger trace. Ranogajec and Geršak also found that EDA finger responses are comparable to EDA foot responses [14].

Poh et al. [13] compared the measurements of mobile EDA sensors on subjects’ wrists against measurements taken with conventional EDA sensors on subjects’ fingers; the stimuli included physical activity, mental arithmetic, Stroop, and horror movies. Then, Poh and other researchers used mobile EDA sensors to collect affective signals from daily activities including reading, walking, and sleeping [13, 16]. They also used mobile EDA sensors in an epilepsy study [12]. The range of applications kept expanding. Chaspari et al. [3] used mobile EDA sensing to model verbal response latencies in autistic children. McDuff et al. [8] used mobile EDA sensing to quantify user emotions in a reflection study of past events. Overwhelmingly, in all these studies researchers used the wristband form of mobile EDA sensing and reported interesting findings.

To summarize, some researchers investigated the strength of EDA responses at several body locations using a conventional EDA device for the measurements; thus, they studied the location effect but not the sensor effect [18]. Other researchers attempted to investigate the sensor effect by using conventional EDA sensors on the subjects’ fingers and mobile EDA sensors on the subjects’ wrists; in doing so, they confounded sensor effect with location effect [13]. Importantly, the stimuli used in all these studies were emotionally complex and thus, non-ideal for inter-subject comparison. This leaves the investigation of mobile EDA sensing incomplete. We aim to fill this gap by answering two questions: (a) What is the EDA sensor effect at body locations with wearability potential? (b) What are the body’s EDA responses in these locations when sympathetic arousal takes place in its most primal form? Stimuli used in prior studies were capable of generating complex emotions, which might depend on the subject’s genotype and phenotype. In contrast, we use auditory startle stimuli in our study. Auditory startle is known to invoke a threshold response on the fingers of healthy subjects [6]. This is the least common denominator when it comes to subject arousal, allowing us to establish a firm basis for comparisons.

An EDA device that successfully captures startle responses on the fingers, meets the ‘gold standard’, as it has the capability of measuring a ubiquitous sympathetic arousal of low intensity and minimal duration at a prime neurophysiological site. Conventional EDA devices belong to this category. Even a gold standard device, however, may not capture startle responses on a different body location (e.g., wrist), where there may be lower concentration of perspiration glands and sparser innervation. Hence, an EDA device should be validated for every intended body location.

Using EDA sensors on body locations that have not been validated against the gold standard may introduce bias into sympathetic measurements in the form of underestimation. With the proliferation of mobile affective studies, this is becoming a pressing question.

**METHODS**

**Ethics Statement.** The experimental procedures were approved by the University of Houston’s Institutional Review Board (IRB), and were performed in accordance with the approved guidelines. Informed consent was obtained from each subject before conducting the experiments.

**Subjects.** We recruited subjects through email solicitations and flyer posting in the university campus community (population about 35,000). We excluded children (< 18), subjects with hearing impairments, and subjects on medications. Age brings psychological and physiological changes that are especially prominent during developmental and late years [2]; for this reason, we did not include children and older adults (> 59) in the subject pool. The use of auditory startle stimuli in the experiment necessitated the exclusion of subjects with hearing problems. Certain medications affect sympathetic responses [5]; to simplify screening and minimize confounding factors, we excluded all medication cases.

**Experimental Design.** After each subject consented, s/he filled out a biographic questionnaire, the Trait Anxiety Inventory (TAI), and the State Anxiety Inventory (SAI) [17]. The last two meant to check if any pathological or extraordinary conditions were biasing responses. Next, we asked each subject to wash her/his hands and feet prior to sensor attachment. We waited 10 min for normal moisture levels to reestablish on their skin, before attaching the sensors. This preparation ensured optimal and uniform skin conditions in the attachment areas for all subjects. We used Galvanic Skin Response probes connected to an ADInstruments PowerLab data acquisition unit (ADInstruments, Bella Vista, Australia) as conventional EDA sensors (C EDA); we used Q Sensor sets (Affectiva, Waltham, Massachusetts) [15] as mobile EDA sensors (M EDA).

This was a validation experiment that aimed to isolate the sensor effect from the location effect in EDA measurements. Ideally, for the sensor validation part we needed to test if the C EDA sensor and the M EDA sensor give the same measurement on the same body location at the same time. The recommended test locations are the fingers and palm, where the sympathetic responses are known to be strong. Because it is impossible to attach the C and M sensors on the same location, however, we used the bilateral symmetry of the human body as a workaround. The results show that the C sensor measurements on the fingers and palm of the right hand were in agreement with the M sensor measurements on the fingers and palm of the left hand, respectively (see RESULTS section). For this to happen both of the following conditions had to be true: (a) The C and M sensors had to have similar capability and (b) the subject pool had to have bilateral symmetry. Hence, our subject pool had bilateral symmetry and this is not surprising, because the great majority of the human population has bilateral symmetry in its functional responses under non-extreme conditions [7].

A total of n = 25 subjects fulfilling the inclusion/exclusion criteria volunteered for the experiment. We attached seven
EDA sensors to each subject: three sets of a conventional EDA sensor (C EDA) on the subject’s right half and four sets of a mobile EDA sensor (M EDA) on the subject’s left half. Specifically, we had the following sensor-location pairs (aka nodes - Fig. 1A):

- Conventional EDA on the right hand fingers (C Fingers)
- Conventional EDA on the right palm (C Palm)
- Conventional EDA on the right wrist (C Wrist)
- Mobile EDA on the left hand fingers (M Fingers)
- Mobile EDA on the left palm (M Palm)
- Mobile EDA on the left wrist (M Wrist)
- Mobile EDA on the left sole (M Sole)

One of the 25 subjects had missing conventional signals due to a technical problem. For another subject the recorded EDA responses across body locations had a low signal to noise ratio, rendering analysis impossible. Hence, we excluded these two subjects from any further consideration. The usable data set included \( n = 23 \) subjects (11 males / 12 females; age: \( 23.91 \pm 8.12 \)). Figure 2, Fig. 3, and Fig. 4 depict the analyzed data set for the finger, wrist, and sole locations, respectively, facilitating cross-sensor and cross-location comparisons.

The main experiment lasted seven minutes for each subject. During these seven minutes, the subject listened to soothing music via a headset in a near dark room. This soothing music was interrupted three times with a glass breaking sound, which served as the auditory startle stimulus. The first stimulus was delivered in the fourth minute. After that, two more stimuli were delivered at one-minute interval each. All sensor measurements were synchronized and recorded throughout the experiment. Hence, our usable data bank accrued 161 EDA signals (23 subjects \( \times \) 7 EDA signals per subject - one for each node).

**Signal Processing.** The C sensor samples at the rate of 25 points per second, while the M sensor samples at the rate of 32 points per second. For this reason, we resampled the M signals down to 25 points per second to establish uniformity across the signal bank. Then, we applied a moving window filter twice to reduce noise. We set the window size at \( W = 125 \) points. We arrived at this selection by performing sensitivity analysis with window sizes \( W = 25 \times k \), \( k = 1, 2, 3, \ldots \) We chose the size of the increment to be 25 points because it matches the minimum resolution (i.e., rate of sampling). At \( k = 5 \) we attained optimal performance by substantially reducing noise without destructing signal information; at \( k > 5 \) the signals exhibited over-smoothing, which affected the performance of the Peak and Onset detection algorithms.

Even under the best circumstances, EDA signals are not only noisy (Fig. 1B) but also variable. An electrodermal response may feature more than one Peak. Such peaks correspond to multiple neurophysiological firings provoked by a single stimulus [2]. Hence, for each stimulus a subject can experience none, one, or several peaks (Fig. 1C). For this reason, signal abstraction is quintessential to fair comparisons; noise filtering...
alone is not sufficient. What is of interest here is the ability of the sensor to measure the essence of the stimulus’ response at the specific site. A normal neurophysiological response (arousal) can be reconstructed to a good approximation from three key points in the corresponding EDA signal [2]: Onset, Peak, and Offset. The Onset point represents the start of the EDA activation; the Peak point represents the culmination of the activation; and, the Offset point represents the ebbing of the activation.

On the noise-filtered signals, we focused on the non-baseline period from $t = 4$ to $7$ min. We divided this period into three segments: $S_1 = [4, 5]$ min, $S_2 = [5, 6]$ min, and $S_3 = [6, 7]$ min; each segment included the respective stimulus $S_1$, $S_2$, $S_3$ and the period up to the delivery of the next stimulus or the end of the experiment. Within each segment we sought arousalss (firings), each characterized by a Peak and an Onset. An Offset might or might not have existed, depending on the recovery rate and the timing of the next arousal. To facilitate curve analysis we performed averaging on the noise-filtered signals reducing the resolution down to two points per second. Then, we executed the following algorithmic steps:

1. **Peak Detection**: We applied a Peak detection algorithm on the non-baseline portion of each signal - $S_1 \cup S_2 \cup S_3$. The template we used for Peak detection was a 5-tuple structure $S(t - 1) \cup S(t + 1) \cup S(t + 2)$, where $S(t)$ is the signal intensity $P$ recorded at time $t$. In essence, this is the discrete definition of the derivative test for locating curve maxima.

Figure 2. $C$ and $M$ EDA responses on the fingers per subject. From the four minutes of baseline only the last minute is depicted in the graphs to economize space. Vertical dotted lines identify stimuli times and inverted triangles denote peaks, facilitating arousal comparison. The $M$ EDA signal for subject S009 was not collected due to technical reasons. The graphs confirm qualitatively the high responsiveness of the fingers location and the agreement between the two sensor modalities.
Figure 3. C and M EDA responses on the wrist per subject. From the four minutes of baseline only the last minute is depicted in the graphs to economize space. Vertical dotted lines identify stimuli times and inverted triangles denote peaks, facilitating arousal comparison. A, Signals manifesting sympathetic responses on the wrist. B, Signals manifesting the absence of sympathetic responses on the wrist (complement to panel A), which were excluded from further processing and analysis. The graphs confirm qualitatively the low responsiveness of the wrist location, a result that was documented quantitatively in the text.

couple of very weak wrist signals, where there is inherently ambiguity as to what constitutes a Peak - see S003 and S013 in Figure 3A.

the finger, wrist, and sole nodes. For the palm node, signals and results were quite similar to the finger node. The Peak detection algorithm accurately located almost all legitimate peaks in our signal bank. It failed to locate peaks only in a
2. Detection of First Onset: Between each stimulus and the first Peak that followed it within the respective segment, we determined the corresponding Onset; this was the Onset of the first arousal in response to the specific stimulus. We computed the intensity of this first Onset as $P_S + 10\% (\text{Peak}_1 - P_S)$, where $P_S$ denotes the signal intensity at the time of the stimulus and $\text{Peak}_1$ denotes the first Peak.

3. Detection of Additional Onsets: When multiple firings took place in response to a stimulus, then more than one Peak points existed within the respective segment $S_i, i = 1, 2, 3$. We located the Onset point corresponding to such an additional Peak in the valley between this additional Peak and the preceding Peak. The template we used for this was a 3-tuple structure fulfilling the following condition: $P(t - 1) > P(t) < P(t + 1)$, where $P(t)$ is the signal intensity $P$ recorded at time $t$.

4. Offset Detection: For each Peak we determined the matching Offset as the 50% × (Peak - Onset) drop-off point. If this point occurred after the next Onset or stimulus, then it was rejected and the Offset was treated as a missing value; this is an indication that the subject had not recovered at the time a new arousal set in.

RESULTS
Peaks are the most characteristic points of an electrodermal response. Hence, our analysis proceeds at two levels:

---

**Figure 4.** M EDA responses on the sole per subject. From the four minutes of baseline only the last minute is depicted in the graphs to economize space. Vertical dotted lines identify stimuli times and inverted triangles denote peaks, facilitating arousal comparison. The graphs confirm qualitatively the responsiveness of the sole location, a result that was documented quantitatively in the text. However, amplitude correlations with the finger and palm locations are not strong, suggesting that precise EDA measurements on the sole may be challenging.
Detection level We pay attention to the occurrence of peaks, as proxies of neurophysiological responses; the absence of peaks signifies the absence of a response at a node.

Measurement level We pay attention to the parameters of the neurophysiological responses recorded at the various nodes. These parameters include the times of occurrence for Onset ($t_{on}$), Peak ($t_p$), and Offset ($t_{off}$) as well as the amplitude of the recorded response ($A = \text{Peak} - \text{Onset}$). The times $t_{on}$, $t_p$, and $t_{off}$ quantify the arousal’s evolution, while the amplitude quantifies the arousal’s intensity (Fig. 1C).

To ensure that the subjects’ trait and state anxiety levels did not affect manifestation of sympathetic activation [9], we obtained their trait (TAI) and state (SAI) anxiety psychometrics [17]. Then, we examined in each node the correlations between the number of peaks detected per subject versus the corresponding TAI and SAI scores. None of these correlations was statistically significant ($p > 0.05$).

Detection Level Analysis
The number of peaks denote responsiveness to stimuli, with zero peaks signifying a non-responsive subject in the specific node. Different subjects exhibited different levels of responsiveness at different nodes. As we observe in Fig. 5, the wrist locations had the highest number of non-responsive subjects irrespective of the sensor type. This indicates that wrists responded poorly to stimuli. The rest of the locations had minimal numbers of non-responsive subjects, indicating the regular presence of arousals in response to stimuli.

We are interested to examine if indeed the proportion of non-responsive subjects differed significantly among the nodes. To do so we use the Binomial distribution. In each node we sampled 23 subjects, where each subject presented a binary outcome: failure, if s/he was non-responsive or success, if s/he had at least one arousal. Hence, for each node if we call $Y_i$ the random variable that denotes the outcome of subject $i$ and $\theta$ the unknown probability of success, we have $Y_i \sim \text{Bernoulli}(\theta)$, giving:

$$Y_i = \begin{cases} 0 & \text{if no peaks in the entire experiment} \\ 1 & \text{if at least one peak in the entire experiment} \end{cases} \quad (1)$$

We count the total number of successes in the 23 subjects, forming the random variable $X|\theta \sim \text{Binomial}(n, \theta)$, $n = 23$:

$$X = \sum_{i=1}^{n} Y_i \quad (2)$$

As we observe in Fig. 5, the maximum likelihood estimates (MLE) at the wrists are quite smaller compared to the other locations. In fact, running a seven-sample test for equality of proportions shows that there are significant differences among the seven nodes ($p < 0.01$). Next, we exclude the two wrist nodes that appear to be the culprit and we compare the proportions in the remaining five nodes. This time the test returns a non-significant number ($p > 0.05$), which indicates that with the exception of the wrist nodes, all the other nodes are statistically equivalent in terms of peak presence.

Measurement Level Analysis
We start the measurement level analysis by studying the relationship of arousal timing between each pair of nodes, taking into account all three stimuli. Figure 6 shows a matrix that is split along its diagonal. The matrix portion below the diagonal shows the scatterplots of Onset ($t_{on}$), Peak ($t_p$), and Offset ($t_{off}$) times for all node pairs. The matrix portion above the diagonal shows the correlation coefficients for the corresponding scatterplots. As indicated both visually and numerically, the timing agreement between nodes is exceptionally high. Hence, the arousal’s evolution is captured accurately at the locations of interest (fingers, palm, wrist, sole), irrespective of the type of sensor used. One has to note, however, that in the case of wrist nodes, the number of points is small. This is consistent with the finding in Fig. 5. Arousal detection at the wrist is rare, but when it occurs, both the C and M sensors track equally well its evolution.

Next, we study the relationship of arousal intensity between each pair of nodes, taking into account all three stimuli. Figure 7 also shows a matrix that is split along its diagonal. The matrix portion below the diagonal shows the amplitude ($A$) scatterplots for all node pairs. The matrix portion above the diagonal shows the correlation coefficients for the corresponding scatterplots. As indicated both visually and numerically, the agreement is poor for pairs involving one wrist node, irrespective of the sensor type attached to this wrist node. Agreement is also poor for pairs involving the mobile sole node. Agreement gets at least moderately strong for pairs involving fingers, palms, or finger-palm, irrespective of the sensing modality. This brings to the fore a problem with EDA sensing: while the timing of the phenomenon is accurately captured across
locations and sensor types, the magnitude of the phenomenon poses a challenge. This challenge is bigger for the sole and insurmountable for the wrist location, irrespective of the sensor type.

**DISCUSSION**

We recognize the importance of quantifiable and objective information the EDA responses can provide to studies of subject affect. Traditionally, EDA sensing is performed on the fingers with conventional EDA devices. Although, this has serious usability problems, it captures well minimal bursts of sympathetic activation. The new mobile EDA sensors, attached on various body locations, have obvious usability advantages. However, one has to be careful not to compromise measurement accuracy or if s/he does, s/he should at least be aware of it.

In this study we found that in response to minimal standardized sympathetic stimuli, conventional EDA devices are in moderate agreement with mobile EDA devices on the fingers and palm. At the same time, we found that both conventional and

---

**Figure 6. Arousal timing relationships between nodes. BELOW THE DIAGONAL:** Scatter-plots of Onset ($t_{on}$), Peak, ($t_p$), and Offset ($t_{off}$) times for the various node pairs. **ABOVE THE DIAGONAL:** The Pearson correlation coefficients that correspond to the strength of the linear relationships depicted in the scatter-plots below the diagonal (all are significant, $p < 0.01$).
Figure 7. Arousal intensity relationships between nodes. BELOW THE DIAGONAL: Scatter-plots of amplitude (A) for the various node pairs. ABOVE THE DIAGONAL: The Pearson correlation coefficients that correspond to the strength of the linear relationships depicted in the scatter-plots below the diagonal (coefficients in bold are significant, \( p < 0.01 \)).

Mobile EDA devices give significantly inferior measurements when attached to the wrist. Actually, not only measurement but also mere detection of sympathetic responses on the wrist is quite challenging. Given the proliferation of mobile EDA devices, often in the form of affective wristbands, and the accompanying marketing hype, these results are a waking call for a more careful examination of operational limitations; otherwise, the introduction of measurement bias in affective studies appears likely. A simple way for researchers to guard against this, is to screen their subjects through a startle experiment, ensuring that they produce sympathetic responses on the wrist before embarking on their study.
Another body location - the sole - that is a candidate site for wearable EDA sensing provides strong detection capability, but relatively poor measurement capability.

This study tested sensors and locations in a stationary context. While this was necessary to minimize confounding factors and establish a clear first-level comparison, it certainly does not account for additional effects that are present in practice. The main such effect is ambulatory motion that needs to be studied in the future.

ACKNOWLEDGMENTS

This material is based upon work supported by the National Science Foundation via grant # IIS-1249208, entitled ‘EAGER: The Effect of Stress and the Role of Computer Mediation on Exam Performance’. Any opinions, findings, and conclusions or recommendations expressed in this paper are those of the authors and do not necessarily reflect the views of the funding agency.

REFERENCES


