Research Article

Linear Regression Analysis Results of the CGM Sensor PPG Comparison Between Predicted PPG Data Using the Candlestick Model and the LEGT Model Against the Measured Sensor PPG Data During a 2-year COVID-19 Quarantine Period for a Type 2 Diabetes Patient Based on GH-Method: Math-Physical Medicine (No. 541)

Gerald C Hsu

EclaireMD Foundation, USA

***Corresponding author**
Gerald C Hsu, EclaireMD Foundation, USA

Abstract
Since 5/5/2018, the author has been applying a continuous glucose monitoring (CGM) sensor device on his upper arm that collected and recorded the complete glucose data continuously at 15-minute time intervals on his iPhone. He accumulated 96 glucoses per day over the past ~3.5 years. As a result, over these 1,272 days, he has compiled a total of 122,112 glucose data and stored them in his database where postprandial plasma glucose (PPG) occupies 45,792 data size and 37.5% of the total glucose database.

During the 2020-2021 COVID-19 quarantine period, he maintained a strict daily routine, without any travel, allowing him to reach an overall healthy lifestyle. Therefore, all of the 19 influential factors of PPG are mainly control by two primary factors: carbs/sugar intake amount (average at 13.1 gram, low-carb diet) and post-meal walking exercise (average of 4,300 steps). These lifestyle improvements helped reduce his PPG waveform amplitudes, including the four associated PPG data of candlestick (aka K-line) model: opening, maximum, minimum, and closing.

Based on the simplified and healthy lifestyle, he can then easily utilize his developed candlestick (aka K-line) model to develop another set of predicted PPG values in addition to the LEGT model results shown in paper No. 540.

In previous research reports, he applied physics concepts and theories, engineering models and equations, mathematical concepts and formulas, computer big data and artificial intelligence (AI) techniques, as well as some statistical approaches. However, the majority of published medical papers he read are mainly based on statistics. As a result, in this article, he selected one of the basic statistical tools, linear regression analysis, to study the comparison between his predicted PPG using K-line model and CGM sensor measured PPG.
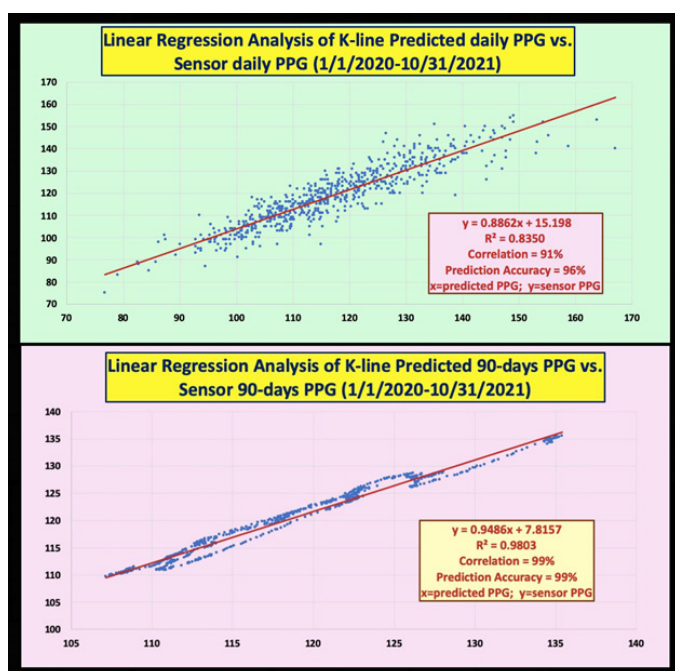
In conclusion, the linear regression analysis results using the K-line model provide more accurate predicted PPG results than using the LEGT model. Actually, both of the maximum PPG and minimum PPG of K-line model are generated through LEGT equation, but the starting PPG of K-line is a measured sensor value instead of the calculation via 0.97*FPG. Moreover, the K-line model contains one extra data point for the closing PPG value. This is why the K-line model has extracted from 4 out of 13 measured sensor PPG data which results in a higher prediction accuracy than the LEGT prediction calculated from carbs/sugar intake grams and post-meal walking steps.

There are three conclusions worth mentioning:

1. From the viewpoint of correlation coefficient in **linear regression analysis** domain, the K-line model's daily calculated PPG vs. sensor measured PPG is 91% and 90-days moving average PPG vs. sensor measured

PPG is 99%. In comparison, the LEGT model's daily calculated PPG vs. sensor measured PPG is 44% and 90-days moving average PPG vs. sensor measured PPG is 65%.

2. From the viewpoint of correlations in **time-domain**, the K-line model's 90-days moving average PPG vs. sensor measured PPG is 99%. In comparison, the LEGT model's 90-days moving average PPG vs. sensor measured PPG is 75%.

3. The "error" or "deviation" of distance between the individual glucose and the red-colored "trend-line" of linear regression model is wider using the LEGT model than the K-line model. This finding is logical to the author from the mathematical viewpoint of his developed math-physical K-line model. Incidentally, the narrower **"error" or "deviation"** of the K-line model's dataset is also directly related to the **higher correlation coefficient R and R square.**

Linear Regression Analysis of K-line Predicted daily PPG vs. Sensor daily PPG (1/1/2020-10/31/2021)

y = 0.8862x + 15.198
R² = 0.8350
Correlation = 91%
Prediction Accuracy = 96%
x=predicted PPG; y=sensor PPG

Linear Regression Analysis of K-line Predicted 90-days PPG vs. Sensor 90-days PPG (1/1/2020-10/31/2021)

y = 0.9486x + 7.8157
R² = 0.9803
Correlation = 99%
Prediction Accuracy = 99%
x=predicted PPG; y=sensor PPG

Introduction

Since 5/5/2018, the author has been applying a continuous glucose monitoring (CGM) sensor device on his upper arm that collected and recorded the complete glucose data continuously at 15-minute time intervals on his iPhone. He accumulated 96 glucoses per day over the past ~3.5 years. As a result, over these 1,272 days, he has compiled a total of 122,112 glucose data and stored them in his database where postprandial plasma glucose (PPG) occupies 45,792 data size and 37.5% of the total glucose database.

During the 2020-2021 COVID-19 quarantine period, he maintained a strict daily routine, without any travel, allowing him to reach an overall healthy lifestyle. Therefore, all of the 19 influential factors of PPG are mainly control by two primary factors: carbs/sugar intake amount (average at 13.1 gram, low-carb diet) and post-meal walking exercise (average of 4,300 steps). These lifestyle improvements helped reduce his PPG waveform amplitudes, including the four associated PPG data of candlestick (aka K-line) model: opening, maximum, minimum, and closing.

Based on the simplified and healthy lifestyle, he can then easily utilize his developed candlestick (aka K-line) model to develop another set of predicted PPG values in addition to the LEGT model results shown in paper No. 540.

In previous research reports, he applied physics concepts and

theories, engineering models and equations, mathematical concepts and formulas, computer big data and artificial intelligence (AI) techniques, as well as some statistical approaches. However, the majority of published medical papers he read are mainly based on statistics. As a result, *in this article, he selected one of the basic statistical tools, linear regression analysis, to study the comparison between his predicted PPG using K-line model and CGM sensor measured PPG.*

Methods
MPM Background
To learn more about his developed GH-Method: math-physical medicine (MPM) methodology, readers can read the following three papers selected from his ~500 published medical papers. The first paper, No. 386 describes his MPM methodology in a general conceptual format. The second paper, No. 387 outlines the history of his personalized diabetes research, various application tools, and the differences between biochemical medicine (BCM) approach versus the MPM approach. The third paper, No. 397 depicts a general flow diagram containing ~10 key MPM research methods and different tools.

In particular, his paper No. 453 illustrates his GH-Method: math-physical medicine in great details, "Using Topology concept of mathematics and Finite Element method of engineering to develop a mathematical model of Metabolism in medicine in order to control various chronic diseases and their complications via overall health conditions improvement".

The Author'S Case of Diabetes and Complications
The author has been a severe type 2 diabetes (T2D) patient since 1996 and weighed 220 lb. (100 kg, BMI 32.5) at that time. By 2010, he still weighed 198 lb. (BMI 29.2) with an average daily glucose of 250 mg/dL (HbA1C of 10%). During that year, his triglycerides reached to 1161 (diabetic retinopathy or DR) and albumin-creatinine ratio (ACR) at 116 (chronic kidney disease or CKD). He also suffered five cardiac episodes within a decade. In 2010, three independent physicians warned him regarding his needs of kidney dialysis treatment and future high risk of dying from severe diabetic complications. Other than cerebrovascular disease (stroke), he has suffered most known diabetic complications, including both macro-vascular and micro-vascular complications.

In 2010, he decided to launch his self-study on endocrinology, diabetes, and food nutrition in order to save his own life. During 2015 and 2016, he developed four prediction models related to

diabetes conditions: weight, PPG, FPG, and A1C. As a result, from using his developed mathematical metabolism index (MI) model in 2014 and the four prediction tools, by end of 2016, his weight was reduced from 220 lbs. (100 kg, BMI 32.5) to 176 lbs. (89 kg, BMI 26.0), waistline from 44 inches (112 cm, nonalcoholic fatty liver disease /NAFLD) to 33 inches (84 cm), average finger glucose reading from 250 mg/dL to 120 mg/dL, and lab-tested A1C from 10% to ~6.5%. One of his major accomplishments is that *he no longer takes any diabetes medications since 12/8/2015.*

In 2017, he has achieved excellent results on all fronts, especially his glucose control. However, during the pre-COVID period of 2018 and 2019, he traveled to approximately 50+ international cities to attend 65+ medical conferences and made ~120 oral presentations. This hectic schedule inflicted damage to his diabetes control, through dinning out frequently, post-meal exercise disruption, jet lag, and along with the overall metabolism impact due to his irregular life patterns through a busy travel schedule; therefore, his glucose control and overall metabolism state were somewhat affected during this two-year heavier traveling period. During 2020 and 2021 with a strict COVID-19 quarantined lifestyle, not only has he written and published ~400 medical papers in 100+ journals, but he has also reached his best health conditions for the past 26 years. By the beginning of 2021, his weight was further reduced to 165 lbs. (BMI 24.4) along with a 6.1% A1C value (daily average glucose at 105 mg/dL), without having any medication interventions or insulin injections. These good results are due to his non-traveling, low-stress, and regular daily life routines. Due to the knowledge of chronic diseases, practical lifestyle management experiences, and his developed various high-tech tools, they contributed to his excellent health status since 1/19/2020, which is the start date of being self-quarantine.

On 5/5/2018, he applied a CGM sensor device on his upper arm and checks glucose measurements every 5 minutes for a total of ~288 times each day. He has maintained the same measurement pattern to present day. In his research work, he uses the CGM sensor glucose at time-interval of 15 minutes (96 data per day). Incidentally, the difference of average sensor glucoses between 5-minute intervals and 15-minute intervals is only *0.4% (average glucose of 114.81 mg/dL for 5-minutes and average glucose of 114.35 mg/dL for 15-minutes with a correlation of 93% between these two sensor glucose curves)* during the period from 2/19/20 to 8/13/21.

Therefore, over the past 11 years, he could study and analyze the collected ~3 million data regarding his health status, medical conditions, and lifestyle details. He applies his knowledge, models, and tools from mathematics, physics, engineering, and computer science to conduct his medical research work. His medical research work is based on the aims of achieving both "high precision" with "quantitative proof" in the medical findings.

The following timetable provides a rough sketch of the emphasis of his medical research during each stage:

- 2000-2013: Self-study diabetes and food nutrition, developing a data collection and analysis software.
- 2014: Develop a mathematical model of metabolism, using engineering modeling and advanced mathematics.
- 2015: Weight & FPG prediction models, using neuroscience.
- 2016: PPG & HbA1C prediction models, utilizing optical physics, AI, and neuroscience.
- 2017: Complications due to macro-vascular research such as cardiovascular disease (CVD), coronary heart disease (CHD) and stroke, using pattern analysis and segmentation analysis.
- 2018: Complications due to micro-vascular research such as chronic kidney disease (CKD), bladder, foot, and eye issues such as diabetic retinopathy (DR).
- 2019: CGM big data analysis, using wave theory, energy theory, frequency domain analysis, quantum mechanics, and AI.
- 2020: Cancer, dementia, longevity, geriatrics, DR, hypothyroidism, diabetic foot, diabetic fungal infection, linkage between metabolism and immunity, and learning about certain infectious diseases such as COVID-19.
- 2021: Applications of linear elastic glucose theory (LEGT) and perturbation theory from quantum mechanics on medical research subjects, such as chronic diseases and their complications, cancer, and dementia. Using metabolism and immunity.it's as the base, he expands his research into cancers, dementia, and COVID-19. In addition, he has also developed a few useful analysis methods and tools for his medical research work.

To date, he has collected nearly 3 million data regarding his medical conditions and lifestyle details. In addition, he has written 536 medical papers and published 500+ articles in 100+ various medical journals, including 7 special editions with selected 20-25 papers for each edition. Moreover, he has given ~120 presentations at ~65 international medical conferences. He has continuously dedicated time and effort on medical research work to share his findings and knowledge with patients worldwide.

Stress, Strain, & Young's Modulus
Prior to his medical research work, he was an engineer in the various fields of structural engineering (aerospace, naval defense, and earthquake engineering), mechanical engineering (nuclear power plant equipment, and computer-aided-design), and electronics engineering (computers, semiconductors, and software robot).

The following excerpts come from the internet public domain, including Google and Wikipedia:

"Strain - ε:
Strain is the "deformation of a solid due to stress" - change in dimension divided by the original value of the dimension - and can be expressed as
ε = dL / L
where
ε = strain (m/m, in/in)
dL = elongation or compression (offset) of object (m, in)

L = length of object (m, in)

Stress - σ:
Stress is force per unit area and can be expressed as
σ = F / A
where
σ = stress (N/m2, lb./in2, psi)
F = applied force (N, lb.)
A = stress area of object (m2, in2)

Stress includes tensile stress, compressible stress, shearing stress, etc.

E, Young's modulus:
It can be expressed as:
E = stress / strain
*   = σ / ε*
*   = (F / A) / (dL / L)*
where
E = Young's Modulus of Elasticity (Pa, N/m2, lb./in2, psi) was named after the 18th-century English physicist Thomas Young.
Elasticity:
Elasticity is a property of an object or material indicating how it will restore it to its original shape after distortion. A spring is an example of an elastic object - when stretched, it exerts a restoring force which tends to bring it back to its original length.
Plasticity:
When the force is going beyond the elastic limit of material, it is into a "plastic" zone which means even when force is removed, the material will not return back to its original state.

Based on various experimental results, the following table lists some of Young's modulus associated with different materials:

Nylon: 2.7 GPa
Concrete: 17-30 GPa
Glass fibers: 72 GPa
Copper: 117 GPa
Steel: 190-215 GPa
Diamond: 1220 GPa

Young's modules in the above table are ranked from soft material (low E) to stiff material (higher E)."

Highlights of Linear Elastic Glucose Theory
Here is the step-by-step explanation for the predicted PPG equation using linear elastic glucose theory as described below:

(1) Baseline PPG equals to 97% of FPG value, or 97% * (weight * GH.f-Modulus).
(2) Baseline PPG plus increased amount of PPG due to food, i.e., plus (carbs/sugar intake amount * GH.p-Modulus).
(3) Baseline PPG plus increased PPG due to food, and then subtracts reduction amount of PPG due to exercise, i.e., minus (post-meal walking k-steps * 5).
(4) The Predicted PPG equals to Baseline PPG plus the food influences, and then subtracts the exercise influences.

The Linear Elastic Glucose Equation is:
*Predicted PPG =(0.97 * GH.f-modulus * Weight) +(GH.p-modulus * Carbs&sugar) - (post-meal walking k-steps * 5)*

Where
(1) Incremental PPG = Predicted PPG - Baseline PPG + Exercise impact
(2) GH.f-modulus = FPG / Weight
(3) GH.p-modulus = Increased PPG, i.e., energy infusion, from Carbs/sugar intake
(4) GH.w-modulus = Decreased PPG, i.e., energy consumption, from Exercise

Therefore, for this case of pre-period's glucose, the modified equation for predicted pre-period's glucose is listed as below:

*Predicted Pre-period's glucose= (FPG * GH.f) + (Carbs/sugar * GH.p) + (walking k-steps * GH.w)*
Where
GH.f = 0.97,
GH.p = 3.22,
GH.w = -5.0

Candlestick (aka K-line) Model
A Japanese merchant, who traded in the rice market in Osaka, Japan, started the candlestick charting around 1850. An American fellow, Steve Nison, brought the candlestick model concept and method to the Western world in 1991. These techniques are largely used in today's stock market to predict the stock price trend.

On 4/17/2018, the author had an idea to study glucose behavior by using the candlestick chart (aka "K-Line") and subsequently developed a customized software to analyze his big data of glucose. The analogies between fluctuations of stock price and glucose value are described as follows:

1. Stock prices are closely related to the psychology of the buyers and sellers, which is similar to the glucoses related to a patient body's biochemical interactions and behavior psychology.
2. Stock price wave of a public traded company is dependent upon its product line, internal management, marketing efforts, and public events and perception. This is remarkably similar to the PPG wave of a diabetes patient being dependent on his/her complex food & diet (buying stock), exercise pattern and amount (selling stock), weather temperature (buying stock), and pancreatic beta cell insulin function (SEC regulations). From a trained mathematician's eyes, both waves are just two similar mathematical representations.
3. When there are more buyers than sellers, the price goes up, which is similar to the glucose value rising when carbs/sugar intake increases (more buyers) or lack of exercise (less sellers).
4. When there are more sellers than buyers, price goes down, which is similar to the glucose value decreasing when carbs/sugar intake decreases (less buyers) or exercise increases

(more sellers).

His standard PPG wave covers 13 data points (every 15 minutes) and 37 data points (every 5 minutes) for a period of 180 minutes, or 3-hours, from the first-bite of his meal. Each PPG waveform contains the following five key characteristic data:

1. "Open" value as his PPG at first-bite, or 0 minute
2. "Close" value as PPG at 180 minutes
3. "Minimum" value as the lowest PPG
4. "Maximum" value as the highest PPG
5. "Average" glucose - average value of 12 recorded PPG data per meal over 3-hours

Based on his meal's candlestick bars, glucose patterns and moving trends can also be observed and analyzed through further mathematical and statistical operations. Finally, he interpreted these operational results with his acquired knowledge of biomedical phenomena of his body in order to discover some hidden medical truth or potential health dangers via TIR analysis.

Since the stock market is much more lucrative than the medical research field, it attracts more talented mathematicians and engineers to work in the highly rewarded financial industry. They even call themselves, "Finance Engineers". On the contrary, most financial rewards in the medical community are distributed to pharmaceutical companies, healthcare institutions, and clinical medical doctors. From the author's personal observation, a large amount of medical research scientists is self-motivated through their interests and dedication, which are mostly associated with either universities or research institutions. They are rarely rewarded financially.

The author is a professionally trained mathematician, physicist, engineer, computer scientist, and a successful entrepreneur. He accidentally wandered into the medical research field due to his strong motivation of saving his own life after suffering many diabetes complications and faced the possibility of death. As a result, he thought about how to import his learned physics principles and theories, mathematical analysis methods, engineering modeling techniques from his academic educations and professional experiences, as well as his accumulated knowledge regarding stock price and other financial analyses techniques, such as the Candlestick model, from his position as the CEO of a public traded corporation, and apply them to his medical research activities. This allowed him to benefit and learn about medicine from using his financial world intellectuals' knowledge along with professional industrial experiences.

Results

Figure 1 shows a combined three 90-days moving average PPG curves with the average PPG values of LEGT predicted PPG at 120.43 mg/dL, K-line predicted PPG at 118.49 mg/dL, and sensor measured PPG at 120.22 mg/dL. *This set of 90-days average PPG values provide 99.8% prediction accuracy using the LEGT model and 98.6% prediction accuracy using the K-line model. It also shows these two equations for the LEGT model and K-line model.*
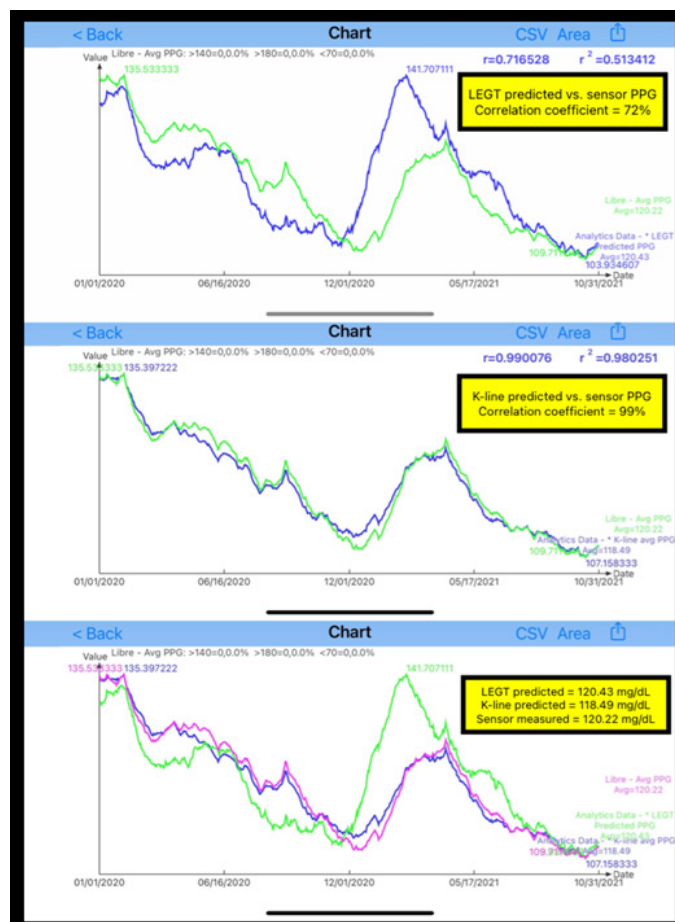

Figure 1: 3 PPG curves and 2 equations of K-line and LEGT

Figure 2 combines three time-domain analysis diagrams of the 90-days moving average PPG together.

The top diagram illustrates the LEGT equation used for calculating his predicted 90-days moving average PPG versus his sensor measured PPG during the COVID-19 quarantine period from 1/1/2020 to 10/31/2021. A moderate high correlation of 75% is observed.

The middle diagram displays the K-line model used for calculating his predicted 90-days moving average PPG versus his sensor measured PPG during the COVID-19 quarantine period from 1/1/2020 to 10/31/2021. An extremely high correlation of 99% is observed.

The bottom diagram combines three 90-days moving average PPG curves together. It should be mentioned that *the correlation between LEGT and K-line model results is 75%.*
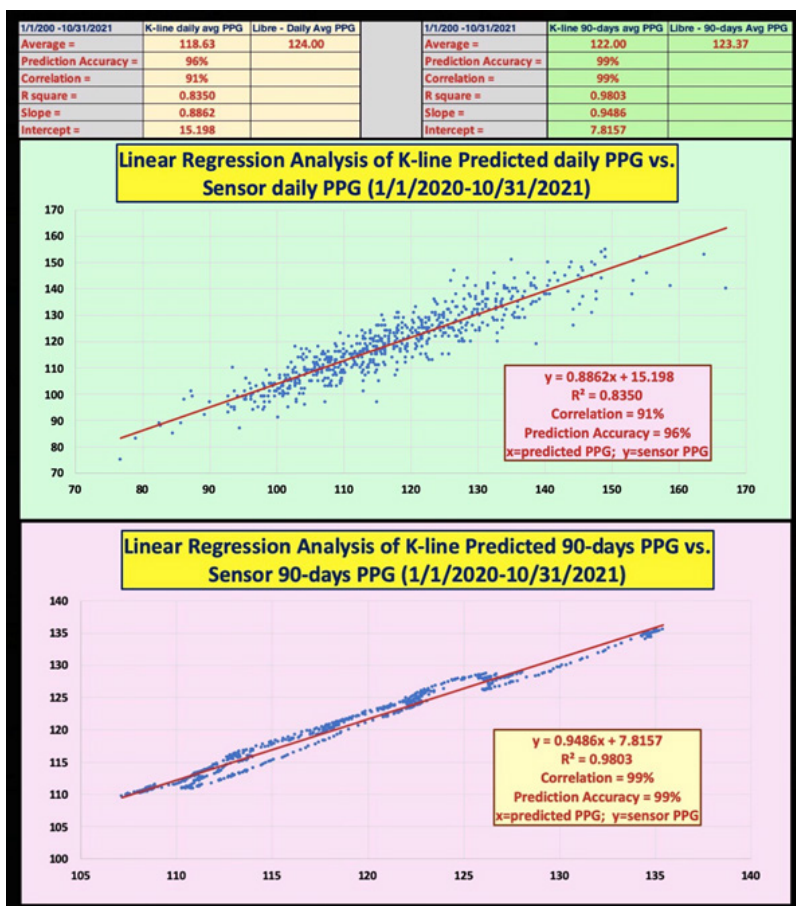

Figure 2: LEGT vs. sensor PPG (top); K-line vs. sensor PPG (middle); 3 PPG curves (bottom)

Figure 3 demonstrates the linear regression analysis results using the daily PPG (top diagram) and 90-days moving average PPG (bottom diagram) of the K-line model.
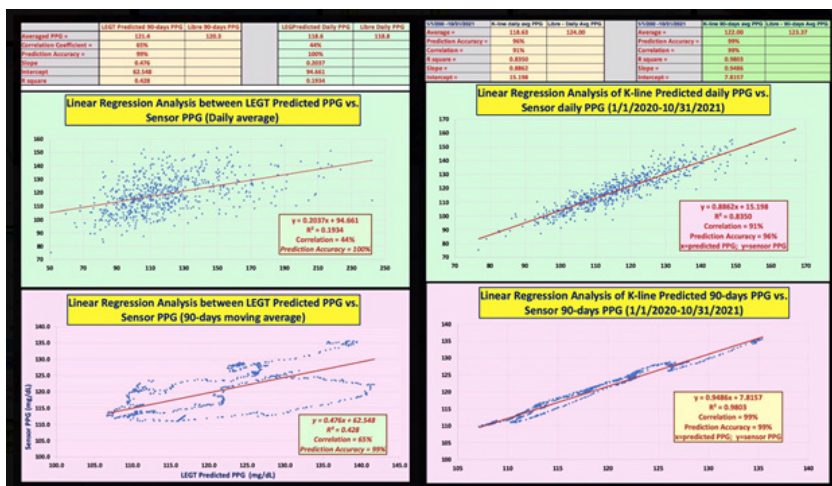

Figure 3: Linear regression analysis of daily PPG (top) and 90-days moving average PPG (bottom)

Conclusive key results are shown below:

K-line Daily PPG
Correlation = 91%
R square = 0.8350
Prediction accuracy = 96%
Slop = 0.8862
Intercept = 15.198
K-line 90-days moving average PPG
Correlation = 99%
R square = 0.9803
Prediction accuracy = 99%
Slop = 0.9486
Intercept = 7.8157

Figure 4 reveals the linear regression analysis results comparison using the K-line model versus LEGT model.

Figure 4: Linear regression analysis results comparison between K-line model and LEGT model

The above-mentioned italics data are K-line PPG model results, while the following data are LEGT PPG model results.

LEGT Daily PPG
Correlation = 44%
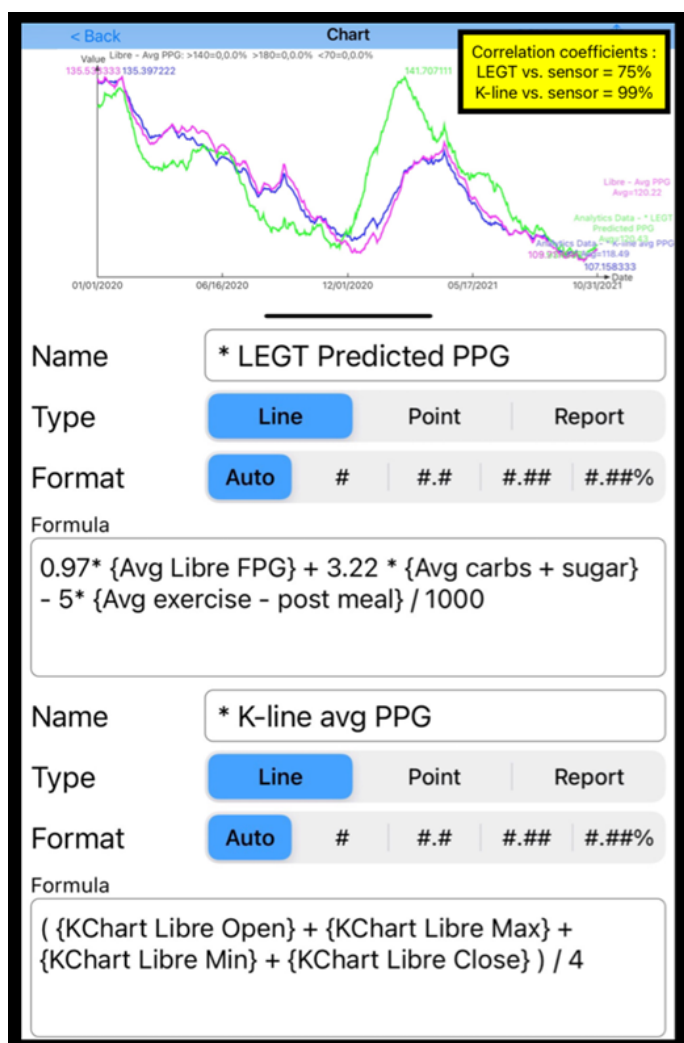*R square = 0.1934*
Prediction accuracy = 100%
Slop = 0.2037
Intercept = 94.661

LEGT 90-days moving average PPG
Correlation = 65%
R square = 0.4280
Prediction accuracy = 99%
Slop = 0.4760
Intercept = 62.548

Based on the displayed datasets and figures, it is obvious that the ***K-line model offers a higher prediction accuracy and detailed characteristics of the linear regression analysis, e.g., correlation and R square. Furthermore, in statistical accuracy, the 90-days moving average PPG results are always better than the daily PPG results.***

Conclusions

In conclusion, the linear regression analysis results using the K-line model provide more accurate predicted PPG results than using the LEGT model. Actually, both of the maximum PPG and minimum PPG of K-line model are generated through LEGT equation, but the starting PPG of K-line is a measured sensor value instead of the calculation via 0.97*FPG. Moreover, the K-line model contains one extra data point for the closing PPG value. This is why the K-line model has extracted from 4 out of 13 measured sensor PPG data which results in a higher prediction accuracy than the LEGT prediction calculated from carbs/sugar intake grams and post-meal walking steps.
There are three conclusions worth mentioning:

1. From the viewpoint of correlation coefficient in linear regression analysis domain, the K-line model's daily calculated PPG vs. sensor measured PPG is 91% and 90-days moving average PPG vs. sensor measured PPG is 99%. In comparison, the LEGT model's daily calculated PPG vs. sensor measured PPG is 44% and 90-days moving average PPG vs. sensor measured PPG is 65%.
2. From the viewpoint of correlations in time-domain, the K-line model's 90-days moving average PPG vs. sensor measured PPG is 99%. In comparison, the LEGT model's 90-days moving average PPG vs. sensor measured PPG is 75%.
3. The "error" or "deviation" of distance between the individual glucose and the red-colored "trend-line" of linear regression model is wider using the LEGT model than the K-line model. This finding is logical to the author from the mathematical viewpoint of his developed math-physical K-line model. Incidentally, the ***narrower "error" or "deviation"*** of the K-line model's dataset is also directly related to the ***higher correlation coefficient R and R square.***

References
For editing purposes, majority of the references in this paper, which are self-references, have been removed for this article. Only references from other authors' published sources remain. The bibliography of the author's original self-references can be viewed at www.eclairemd.com.

Readers may use this article as long as the work is properly cited, and their use is educational and not for profit, and the author's original work is not altered.