

Co-integration for Soft Commodities with Non Constant Volatility

E. Chanol, O. Collet, N. Kostyuchyk, T. Mesbah, and Quoc Hoang Long Nguyen

Abstract—In this paper, a pricing model is proposed for co-integrated commodities extending Larsson model. The futures formulae have been derived considering a linear combination of a Brownian motion and an Ornstein-Uhlenbeck process describing the co-integration relationship of the different futures prices commodities. Tests have been performed with a non-constant volatility in order to fit better the real behavior of the volatility. The model has been applied to energy commodities (gas, CO₂, energy) and soft commodities (corn, wheat). Results show that first, the model can be used with different kind of commodities at the cost of a proper parameters calibration and in second, using a non-constant volatility leads to more accurate short term prices, which provides better evaluation of Value-at-Risk and more generally improves the risk management.

Index Terms—Co-integration, risk management, soft commodities, value-at-risk.

I. INTRODUCTION

In the trading of raw materials, an asset value is often strongly linked to another one. For example, electricity price depends on the coal used to produce it.

The study of dependences and relations between various raw materials is therefore essential, and co-integration is nowadays a common tool for investigating dependences in multivariate time series [1]. This corresponds to existence of a relationship between the prices which makes a linear combination of them stationary, even though individually they are non-stationary.

In Ref. [2], multiple energy spot prices such as gas and power between different markets are considered. A co-integrating multi-market model framework is developed by connecting different single-market spot-price models and it is shown that gas prices are strongly co-integrated with power prices for a specific market. The different commodities must therefore be taken in the same market.

In Ref. [3] the evolution of electricity and gas prices from the UK market is modeled and option price on the spark spread is simulated. Despite using co-integration technique, the series are considered as normal inverse Gaussian processes and copula are calculated which can strongly influence the option prices. The considered energy commodities as well as a different method for linking both of them are interesting for comparing the model with a co-

integration model.

In Ref. [4] the price relationship of primary agricultural commodities, exchange rates and oil prices are examined, as nowadays many agricultural commodities can be used to produce energy. Analysis of co-integration vector of the different prices over time showed that commodity prices are linked to oil for corn, cotton and soybeans. In the model, the question of seasonality effect is raised, which can be interesting to consider.

In Ref. [5] a continuous-time co-integrated asset dynamics using spot prices of crude oil and gasoline is proposed. Results showed that both commodities are co-integrated and an interesting calibration method for commodity markets is given. However, most results focus on spot price modeling and then deduce the future prices endogenously.

Present work is based on [2], where it is shown that ignoring co-integration and using simple correlation based price models heavily overestimates the spread variation over long time horizons. This can lead to drastically overestimate Value-at-Risk and excessive capital requirements. Here to analyze this point, two main objectives are looked for: first one is to check improvement nature in Larsson-Nossman model [6] by when specifying non-constant volatility into the parameters. The second one is to apply this model to soft commodities in order to show that the model is generic [7]. Soft commodities can be defined as commodities which can be grown such as coffee, cocoa, sugar, corn, wheat, soy bean and fruits [8]. Adding a non-constant volatility makes the model more realistic, and a reduction in price oscillations can be observed. Reducing uncertainty in this type of model is very important for many market actors. Indeed, a significant decrease on price variation is observed during the first hundred days. Present analysis is meant for pricing of raw materials, and many producers and consumers are sensitive to such variations. They can partly be explained by factors like seasonality, as soft commodities require time to be produced and as most of the time, prices are negotiated before the hopping (forward contracts). The inherent gap between supply and demand causes a constraint on the short-term supply and instabilities on agricultural markets. In a same way, markets can be affected by factors such as weather conditions and diseases affecting crops. All these elements contribute to the complexity of soft commodities market, and justify why modeling such market represent a big challenge with large economic impact.

II. MODEL

In this section, the main equations for calculating a commodity future price are stated. From [6], the futures

Manuscript received July 25, 2014; revised January 10, 2015.

The authors are with the Financial Engineering Department, ECE – Paris, Paris, 75015 France (e-mail: nguyen.ql@gmail.com, kostyuch.ua@gmail.com, tarek_mesbah001@gmail.com, chanol.elisabeth@gmail.com).

prices of N commodities at time t with delivery time T , are given by $\Gamma(t; T)(\cdot) = \exp[\mathbf{F}(t; T)(\cdot)]$ with $\Gamma = \text{col}[\Gamma_1, \Gamma_2, \dots, \Gamma_N]$, $\mathbf{F} = \text{col}[F_1, F_2, \dots, F_N]$. The log futures prices $\mathbf{F}(t, T)$ obey the equations :

$$d\mathbf{F}(t; T) = \boldsymbol{\kappa}Z(t)dt + \boldsymbol{\sigma}(t; T).d\mathbf{W}(t) \quad (1)$$

with $\boldsymbol{\kappa} = \text{col}[\kappa_1, \kappa_2, \dots, \kappa_N]$, $\boldsymbol{\sigma} = \text{diag}[\sigma_1, \sigma_2, \dots, \sigma_N]$ and $\mathbf{W} = \text{col}[W_1, W_2, \dots, W_N]$, and where $Z(t) = \beta_0 + \boldsymbol{\beta}^T \mathbf{F}(t, T)$ with $\boldsymbol{\beta} = \text{col}[\beta_1, \beta_2, \dots, \beta_N]$. From (1) and using Ito formula one gets

$$d[Z(t)\exp(-Bt)] = \boldsymbol{\beta}^T \boldsymbol{\sigma}(t; T). \exp(-Bt) d\mathbf{W}(t) \quad (2)$$

where $B = \boldsymbol{\beta}^T \boldsymbol{\kappa}$, and integration of (2) gives

$$Z(t) = Z(0)\exp(Bt) + \int_0^t \boldsymbol{\beta}^T \boldsymbol{\sigma}(s; T). \exp(B(t-s)) d\mathbf{W}(s) \quad (3)$$

Reporting in (1) one finally obtains its solution in the form

$$\mathbf{F}(t, T) = \mathbf{F}(0, T) + Z(0)E(t)\boldsymbol{\kappa} + \left\{ \int_0^t \boldsymbol{\beta}^T \boldsymbol{\sigma}(s; T). E(B, t-s) d\mathbf{W}(s) \right\} \boldsymbol{\kappa} + \int_0^t \boldsymbol{\sigma}(s; T) d\mathbf{W}(s) \quad (4)$$

with the function $E(B, u) = B^{-1}[\exp Bu - 1]$.

To completely determine forward price vector $\mathbf{F}(t, T)$ time dependence of volatilities has to be fixed. On simple grounds volatility is expected to be small at the beginning and to increase with time with highest value at maturity. This is relevant because volatility can be understood as uncertainty, and making projections on a long period of time is more subject to uncertainty than on a short one. Here one will take for each diagonal element the semi-empirical form

$$\sigma_k(u; T) = \theta_k \exp[-\alpha_k(T-u)] \quad (5)$$

The exponential part reduces volatility when the time is small whereas it tends to 1 when the time gets close to maturity. Aside ease in integral computation, exponential dependence in (5) is reminiscent of previous results in LQR optimal control [9], where transpose system with opposite sign eigenvalues is associated to actual system for guiding system trajectory toward the prescribed target. It is also in agreement with fading away decay of fluctuating functions when averaged. With this dependence, forward price is given by

$$\mathbf{F}(t, T) = \mathbf{F}(0, T) + Z(0)E(B, t)\boldsymbol{\kappa} + B^{-1} \left\{ \boldsymbol{\beta}^T \boldsymbol{\sigma}(0, T). [\mathbf{Y} - \exp(-Bt)(\mathbf{B}\mathbf{X} + \mathbf{Y})] \right\} \boldsymbol{\kappa} + \boldsymbol{\sigma}(0, T). \mathbf{Y} \quad (4)$$

where N -vectors \mathbf{X} and \mathbf{Y} are containing all randomness (i.e. the risk) associated respectively with the commodities and co-integration dynamics in the form

$$\begin{aligned} \mathbf{Y}(t) &= \int_0^t \exp(\boldsymbol{\alpha}u) d\mathbf{W}(u); \\ \mathbf{X}(t) &= \int_0^t \exp(\boldsymbol{\alpha}u) E(B, u) d\mathbf{W}(u) \end{aligned} \quad (5)$$

Introducing component wise standard deviation v_X and v_Y of \mathbf{X} and \mathbf{Y} , the following normalized variables $\bar{\mathbf{X}}$ and $\bar{\mathbf{Y}}$ will

be introduced by $\mathbf{X} = v_X \bar{\mathbf{X}}$ and $\mathbf{Y} = v_Y \bar{\mathbf{Y}}$ so that $\bar{\mathbf{X}} \sim N(0, 1)$, $\bar{\mathbf{Y}} \sim N(0, 1)$. As these processes are not independent, they can be represented in the form

$$\bar{\mathbf{Z}} = \text{col}[\bar{\mathbf{Y}}, \bar{\mathbf{X}}] = \mathbf{H}(t). \boldsymbol{\mathcal{E}}(t, \cdot) \quad (6)$$

where $\boldsymbol{\mathcal{E}}(t, \cdot) = \text{col}[\varepsilon_1(\tau, \cdot), \varepsilon_2(\tau, \cdot), \dots, \varepsilon_N(\tau, \cdot)]$ is a N -vector of independent random variables with a distribution $N(0, 1)$ and $\mathbf{H}(t) = \text{Tab}[\eta_{i,j}(t)]$ a lower triangular matrix with $\eta_{i,j}(t) = 0$ for $j > i$, and normalization $\sum_j (\eta_{i,j})^2 = 1$, the coefficients of which are determined by calculation of covariance between the various components of $\bar{\mathbf{Z}}$. First one gets from (5)

$$\begin{aligned} \text{Cov}[Y_i, Y_j] &= \rho_{ij} E(\alpha_i + \alpha_j, t); \\ \text{Cov}[Y_i, X_j] &= B^{-1} \rho_{ij} [E(\alpha_i + \alpha_j + B, t) - E(\alpha_i + \alpha_j, t)]; \\ \text{Cov}[X_i, X_j] &= B^{-1} \rho_{ij} [E(\alpha_i + \alpha_j + 2B, t) - 2E(\alpha_i + \alpha_j + B, t) + E(\alpha_i + \alpha_j, t)] \end{aligned} \quad (7)$$

and standard deviation $(v_{X_i})^2 = \text{Cov}[X_i, X_i]$, $(v_{Y_i})^2 = \text{Cov}[Y_i, Y_i]$ with $\rho_{ij} = \langle dW_i, dW_j \rangle$ the elements of correlation matrix between random fluctuations acting on system (1). Using (6, 7) and definition of $\bar{\mathbf{Y}}, \bar{\mathbf{X}}$ one similarly obtains

$$\text{Cov}[\bar{Z}_i, \bar{Z}_j] = \sum_{k \leq \min(i, j)} [\eta_{i,k} \eta_{j,k}] = (v_{Z_i} v_{Z_j})^{-1} \text{Cov}[Z_i, Z_j] \quad (8)$$

from which all coefficients $\eta_{i,k}(t)$ can be easily evaluated.

From these coefficients, it is possible to simulate the considered processes and therefore to make a pricing. With the prices, the important Value-at-Risk (VaR) in risk management can be derived. The VaR is the maximal loss that a financial product holder can make. The product can be equities or derivatives. Mathematically, the VaR is the maximal loss value with in general a 95% confidence probability. The VaR computation determines the amount to save in order to hedge the market risk (to make the risk neutral). The smaller is the VaR, the less assets the company has to save, allowing more opportunities for investments or savings. To calculate the VaR N simulations are performed with the model giving daily prices between the spot today and the maturity at last considered day. For each simulation, the difference between generated prices and spot prices is evaluated and the difference is stored in a matrix. At the end of the N simulations, the minimal value of all the differences is computed. These minima are stored as a vector sorted in ascent order, and starting from the last vector, the value corresponding to 5% of the total number of vectors is selected. This value is the VaR. the precision of which is increased with N .

III. SIMULATIONS

In this section, simulations are performed with three commodities (energy, gaz and CO₂) used in [10] in order to compare their results with the proposed time dependent specification of volatility. The motivation of a new volatility is simple. First considered Black-Scholes model is a mathematical model focused on derivatives giving theoretical price estimate of European-style options. Nowadays, despite its age and simplicity, this model still has an application, showing it may sometimes be more

interesting to work with model simplicity than trying to improve its complexity by adding non-constant or stochastic parameters. Present aim is to improve Larsson-Nossman model while preserving its simplicity. Therefore, constant volatility (a weak point in the model) has been modified into time dependent one in (5).

Aside the fact that the chosen volatility depicts better markets behavior than a constant one, it does not add complication while providing better pricing results. Indeed, the volatility exponential part, depending on time and maturity, is easy to implement in the model and does not make the covariance computations harder. To be relevant, tests are made under the same conditions meaning that the same parameters are considered, and the spread of these commodities will be calculated with constant and non-constant volatility.

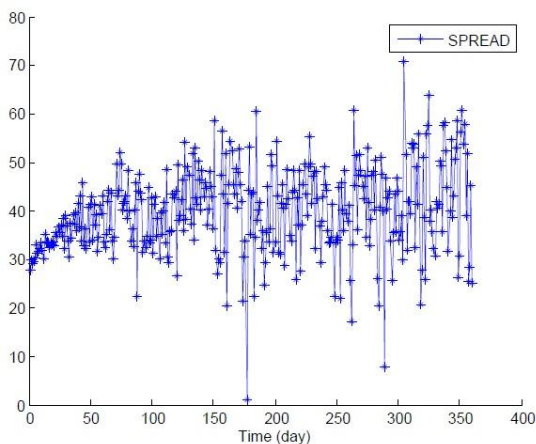


Fig. 1. Simulation of the spread prices of 3 commodities (gas, CO₂, energy) using larsson-nossman model with constant volatility and same parameters.

The curves in Fig. 1 oscillate strongly and quickly in time, while the amplitude of oscillations is smaller with non-constant volatility. The oscillations start to increase around 150 days whereas oscillations become large already around 40 days with constant volatility, implying that short term expectations are more accurate with non-constant volatility.

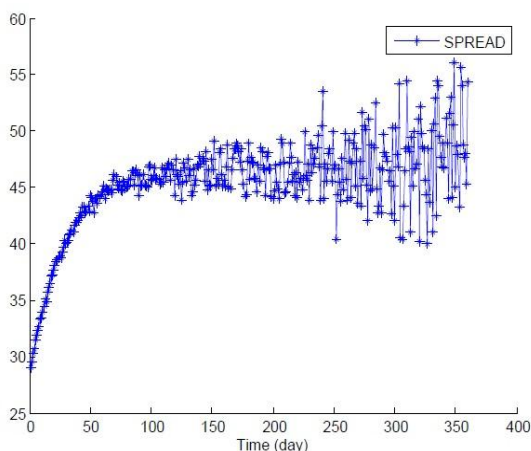


Fig. 2 . Simulation of the prices spread of 3 commodities (gas, CO₂, energy) using larsson-nossman model with non constant volatility and same parameters.

The VaR can be derived from simulated prices. In order to make comparisons with Larsson and Nossman tests, gas price VaR has been simulated. The VaR is calculated for different maturities with 3 months interval and has been

plotted for the different maturities, with a maximal maturity of 36 months (3years).

We have computed VaR of gas price with Larsson-Nossman model using the constant volatility and the non-constant volatility. With the constant volatility, it can be observed that the larger is the maturity the bigger is the loss risk, with value 15 reached at maturity of 3 years. However with the non-constant volatility, the VaR is smaller for the same considered maturity with a maximal value reached for the maturity of 3 years and a loss of 5. It should be noted that both calculations are made with same parameters, except for the parameters which have been chosen in coherence with the model. As expected, using an exponential part in volatility makes the uncertainty smaller at short term periods. The Figures illustrate this phenomenon with better pricing results thanks to this new volatility.

It is interesting at this point to test theoretical present model with real data market. Energetic data prices would be interesting as Larsson-Nossman model has been initially designed for this type of commodities, but the only available real data prices are concerning soft commodities from Euronext market: wheat and corn prices from 02/01/2013 to 23/09/2013. On the other hand, it is largely expectable that there exist strong correlations between different segments in an integrated economy [11], so it is interesting to apply present theoretical study to such domains as well. As present study focuses on co-integrated commodities, it is first required to check this property for selected soft ones. For this purpose, augmented Dickey-Fuller (ADF) test [12]-[21] with the null hypothesis that the time series is integrated of order one, and Engle-Granger test [1] are both made in order to determine whether both commodities are co-integrated. Next step is to calibrate model parameters, ie to determine best values giving results closest as possible to real data prices. As calibration process is fundamentally an optimization problem, the parameters are determined for each prices, meaning there are as many parameter sets as there are input prices. Here as the model only requires one single set for simulating the results, average of all parameters is considered.

TABLE I: OPTIMAL VALUES OF MODEL PARAMETERS, USING EURONEXT SOFT COMMODITIES

Parameters	Optimal value considering soft commodities sample
K1	-0.0756
K2	0
Beta0	2.4144
Beta1	0.8929
Beta2	0.4390
Lambda1	1
Lambda2	1
Rho	-13.4710
Rho12	0.6342
Alpha1	5.0191
Alpha2	3.0036

On Fig. 3, it can be observed that simulated prices fit quite well to data market prices, strongly indicating that the model can be used for other commodities than energetic ones. However agricultural markets are particularly complex and their soft commodities are not always co-integrable such as energetic raw materials. This is why co-integration tests must be performed before using proposed theoretical model.

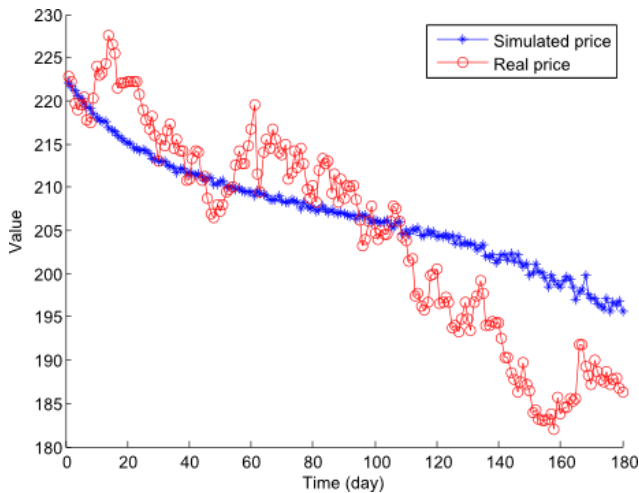


Fig. 3. Comparison between simulated and real data market prices for wheat.

IV. CONCLUSION AND PERSPECTIVE

In conclusion, considering a non-constant volatility definitely improves the results of Larsson-Nossman model because uncertainty relative to time affects the prices with a delay of 100 days approximately, making the evaluation over the 100 beginning days more accurate. Moreover, the less oscillating are the prices, the smaller is the VaR which is an important aspect in risk management. So the use of proposed non-constant volatility brings notable improvements, to the extent that parameters are correctly calibrated, especially the parameter α affecting volatility.

While making the results better, a unique solution depending on entry data is proposed. Mathematical study shows that the generic model is analytically solvable and independent of parameter choice. Present model is not perfect however and can be made slightly better by taking both present non-constant volatility and soft commodities into account. Introducing parameter optimization could also be very interesting to improve model accuracy and produce smaller pricing difference with real market prices. This includes study of parameter sensitivity for making an analogy with the Greeks (delta, gamma, vega).

ACKNOWLEDGMENT

The authors are very much indebted to ECE Paris Graduate School of Engineering for having provided the environment where the work has been developed, to Dr Y. Rakoto for guidance and Pr. M. Cotsaftis for help in preparation of the manuscript.

REFERENCES

- [1] R. F. Engle and C. W. J. Granger, "Co-integration and error correction: Representation, estimation and testing," *Econometrica*, vol. 55, pp. 251-276, 1987.
- [2] C. de Jong and S. Schneider, "Co-integration between gas and power spot prices," *The Journal of Energy Markets*, vol. 2, no. 3, pp. 27-46, 2009.
- [3] F. E. Benth and P. C. Kettler, "Dynamic copula models for the spark spread," *Quantitative Finance*, vol. 11, no. 3, pp. 407-421, 2011.
- [4] A. Harri, L. Nalley, and D. Hudson, "The relationship between oil, exchange rates, and commodity prices," *J. Agricultural and Applied Economics*, vol. 41, no. 2, pp. 501-510, 2009.
- [5] M. C. Chiu and H. Y. Wong, "Mean-variance portfolio selection of co-integrated assets," *J. Economic Dynamics and Control*, vol. 35, pp. 1369-1385, 2011.

- [6] E. Schwartz and J. E. Smith, "Short-term variations and long-term dynamic in commodity prices," *Management Science*, vol. 46, no. 7, pp. 893-911, 2000.
- [7] J. A. Frankel and A. K. Rose, "Determinants of agricultural and mineral commodity prices," HKS Faculty Research Working Paper Series RWP10-038, John F. Kennedy School of Government, Harvard University, July 28, 2009.
- [8] H. Geman, *Commodities and Commodity Derivatives: Modeling and Pricing for Agricultural, Metals, and Energy*, Wiley Finance, West Sussex, 2005.
- [9] J. C. Hsu and A. U. Mayerk, *Modern Control Principles and Applications*, McHill, New York, 1968.
- [10] K. Larsson and M. Nossman, "Risk management of co-integrated commodities," Working Paper, SSRN eLibrary, 2011.
- [11] D. A. Dickey and W. A. Fuller, "Distribution of the estimators for autoregressive time series with a unit root," *J. Amer. Statistical Association*, vol. 74, pp. 427-431, 1979.
- [12] J. Campiche, H. Bryant, J. Richardson, and J. Outlaw, "Examining the evolving correspondence between petroleum prices and agricultural commodity prices," in *Proc. Amer. Agricultural Economics Assoc. Annual Meeting*, Portland, OR, July 29-August 1, 2007.
- [13] J. C. Duan and S. R. Pliska, "Option valuation with co-integrated asset prices," *J. Economic Dynamics and Control*, vol. 28, pp. 727-754, 2004.
- [14] M. C. Chiu and H. Y. Wong, "Co-integration and stochastic correlation models for commodity derivatives," in *Proc. APAD 2013*, Busan, Korea, Aug. 22-23, 2013.
- [15] H. Askari and N. Khrichene, "Oil price dynamics (2002-2006)," *Energy Economics*, vol. 30, pp. 2134-2153, 2008.
- [16] S. Lardica and V. Mignona, "Oil Prices and economic activity: An asymmetric co-integration approach," *Energy Economics*, vol. 30, pp. 847-855, 2008.
- [17] S. Westgaard, M. Estenstada, M. Seima, and S. Frydenberga, "Co-integration of ICE gas oil and crude oil futures?" *Energy Economics*, vol. 33, pp. 311-320, 2011.
- [18] M. A. H. Dempster, E. Medova, and K. Tang, "Long term spread option valuation and hedging," *J. Banking and Finance*, vol. 32, pp. 2430-2540, 2008.
- [19] T. H. Yu, D. A. Bessler, and S. Fuller, "Co-integration and causality analysis of world vegetable oil and crude oil prices," in *Proc. Amer. Agricultural Economics Assoc. Annual Meeting*, Long Beach, CA, July 23-26, 2006.
- [20] Q. Zhang and M. Reed, "Examining the impact of the world crude oil price on China's agricultural commodity prices: The case of corn, soybean, and pork," in *Proc. South Agricultural Economics Assoc. Annual Meeting*, Dallas, TX, February 2-5, 2008.
- [21] K. Hanson, S. Robinson, and G. Schluter, "Sectoral effects of a world oil price shock: Economy-wide linkages to the agricultural sector," *J. Agricultural and Resource Economics*, vol. 18, no. 1, pp. 96-116, 1993.



E. Chanol was born in Schoelcher, Martinique on April 16, 1992. She got a baccalaureat in sciences in the high school S éminaire Coll ège Sainte-Marie, Fort de France (97200), Martinique, and after she did classes Pr éparatoires aux Grandes Ecoles in maths-physics specialization in the High School Bellevue, Fort de France (97200). She is currently an undergraduate student in financial engineering (9th semester) at the Engineering School ECE Paris, Paris, (75015) France.

She did her first work as an assistant of the marketing director and of the assistant director general for a month in HD Assurances, a 6 weeks internship as home-officer at BNP Paribas, and 3 months internship as an assistant project management in the federation of tennis.



O. Collet was born in Auray, Bretagne (56550), France on February 8, 1991. After getting a baccalaureat in sciences in the High School Sainte-Anne d'Auray, Auray (56400), France, she did classes Pr éparatoires aux Grandes Ecoles in maths physics specialization (4th semester) in the High School Clemenceau, Nantes, France, (44000). She is currently an undergraduate student in Financial Engineering (9th semester) at the engineering school ECE Paris, Paris, (75015) France.

She did her first work as a logistician for 3 weeks in a pharmacy, a 6 weeks internship as a business function support in juridical department of the company Saint-Gobain, and 4 months internship as an assistant project management in the company Natixis, 30 Avenue Pierre Mend ès France,

Paris, (75013), France. She also gives Maths and Physics lectures to high school level students.



N. Kostyuchykh was born in Kremenets, Ukraine on April 19, 1991. He came studying in France when he was 13 years old in 2004. After getting a baccalaureat in science in the High School Gerson, Paris (75116), France, he did classes Préparatoires aux Grandes Ecoles in maths-physics specialization (4th semester) in the high school Charlemagne, Paris, France (75004).

He is currently an undergraduate student in Financial Engineering (9th semester) at the engineering school ECE Paris, Paris (75015) France. He did his first work as a technician for 2 months in “La Clinique de Mousseau”, a 2 months internship as an IT-developper in the company Mita-Teknik and 4 months internship as a JAVA developper in the company Koltech. He also gives maths and physics lectures to high school level students.



Tar ek Mesbah was born in Paris 16eme, France on April 28, 1991. He then studied at Saint-James high school in Neuilly sur Seine and obtained his Baccalaureat with merits. He is currently an undergraduate at ECE, Paris in the Quantitative Finance Department.

He made a three months internship as a mobile developer in a start-up company specialized in information systems.



Quoc Hoang Long Nguyen was born in Créteil, Val de Marne (94), France on November 1, 1991. After getting a baccalaureat in sciences in the High School Louise Michel, Champigny sur Marne (94500), he did classes Préparatoires aux Grandes Ecoles in physics-chemistry specialization (2nd semester) in the High School Marcelin Berthelot, Saint-Maur (94), France. He is currently an undergraduate student in financial engineering (9th semester) at the engineering school ECE Paris, Paris, (75015) France.

He did his first work as a logistician for 3 months and 3 months internship as an IT-developper both for the company Kaufman&Broad, 127 avenue Charles de Gaulle, Paris, France.