

Supplementary Information: Regression resistivity model

Resistivity mapping of SiC wafers by quantified Raman spectroscopy

Elisa Calà ^a, Simone Cerruti ^c, Cristina Sanna ^b, Marco Maffè^b, Wen Chin Hsu^d, Man Hsuan Lin^d, Luciano Ramello^a and Giorgio Gatti* ^a

a. Department for Sustainable Development and Ecological Transition, University of Eastern Piedmont 'Amedeo Avogadro', Piazza Sant' Eusebio 5, 13100 Vercelli (Eastern Piedmont 'Amedeo Avogadro', Piazza Sant'Eusebio 5, 13100 Vercelli (Italy)

b. MEMC Electronic Materials SpA – GlobalWafers Co., Viale Gherzi 31, 28100 Novara (Italy)

c. Department of Sciences and Technological Innovation, University of Eastern Piedmont 'Amedeo Avogadro', Viale Teresa Michel 11, 15121 Alessandria (Italy)

d. GlobalWafers Co., Ltd- No.8, Industry E. Rd. II, Hsinchu Science Park, Hsinchu 30075 (Taiwan)

† Footnotes relating to the title and/or authors should appear here.

Electronic Supplementary Information (ESI) available: [details of any supplementary information available should be included here]. See DOI: 10.1039/x0xx00000x

We created a linear regression model to correlate resistivity by ADE and four variables of the Raman spectrum (intensity and FWHM of the bands called P5 and P6 in the LOPC region): the coefficient of determination R^2 is ca. 0.96. This model is written using the *R* language, a free software environment for statistical computing and graphics, and the package *caret*, a framework specifically made to train and test statistical models based on various techniques. It should be noted that, before training the model itself, data were autoscaled (consisting of the subtraction of each variable of the respective mean value and then division by its standard deviation, so that they are all on a common scale).

The model was stored in a .RDS (R Data Serialization) file called *Model SiC resistivity.RDS*: below is reported the code which can be used to import it, together with the output of the summary method showing the most relevant information about it (properties of residuals' distribution, values of the estimated coefficients and some statistical indexes about its fitting ability).

```
require(caret)
RES_model <- readRDS("Model SiC resistivity.RDS")
summary(RES_model)
##
## Call:
## lm(formula = .outcome ~ .. data = dat)
##
## Residuals:
## Min      1Q      Median      3Q      Max
## -1.5842  -0.5796   0.2014    0.4377   1.5812
##
## Coefficients:
##              Estimate Std. Error t-value Pr(>|t|)
## (Intercept)  18.4664   0.1153   160.223 < 2e-16 ***
## Height_P5     2.0488   0.2201    9.310 1.86e-11 ***
## Height_P6     1.2101   0.2126    5.692 1.39e-06 ***
## FWHM_P5       2.0567   0.6771    3.037 0.004242 **
## FWHM_P6      -2.8521   0.6882   -4.144 0.000177 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.7645 on 39 degrees of freedom
## Multiple R-squared:  0.9603.    Adjusted R-squared:  0.9562
## F-statistic: 235.9 on 4 and 39 DF.  p-value: < 2.2e-16
```