Selecting varistor clamping voltage: Lower is not better!

François D. Martzloff and Thomas F. Leedy
National Institute of Standards and Technology

Reprinted with permission from *Proceedings, 1989 Zurich Symposium*

Significance:
Part 7: Mitigation techniques

The early applications of zinc oxide varistors, quickly renamed metal-oxide varistors (MOV) in the field of low-voltage AC power circuits were driven by the misconception (seen as such with hindsight) that the lower their clamping voltage, the better the surge protection afforded to the equipment they were intended to protect. This perception was further enhanced by the publication of an Underwriters Laboratory Safety Standard on “Transient Voltage Surge Suppressors” which borrowed a tabulation from an IEC document listing voltage ratings for equipment.

Unfortunately it was not recognized that the 330 V level selected as the lowest row of that IEC table was intended for circuits up to only 50 V, not the 120 V of the consumer electronics equipment. Marketing pressure driven by the “Lower is Better” syndrome prevented a belated recognition of the disadvantages of such low clamping voltage. From the commercial-free platform of NIST, this paper, with some hindsight (work performed in the late eighties) was an attempt to call attention to the problem. In spite of this tutorial effort, the situation was not corrected and to this day, some commercial offerings still rate their clamping voltage as 330 V.

Reprinted from the proceedings of the
1989 Zürich International EMC Symposium

SELECTING VARISTOR CLAMPING VOLTAGE: LOWER IS NOT BETTER !

François D. Martzloff and Thomas F. Leedy
National Institute of Standards and Technology
Electrosystems Division
Gaithersburg, MD 20899

ABSTRACT - Surge protective devices, such as varistors, are applied to protect sensitive load equipment against power-line surges. The need to provide low clamping voltage for protection of equipment with low inherent immunity must be balanced against the risk of premature aging of the protective device. Lower clamping voltage causes more frequent interventions of the protective device, accelerating its aging. The paper describes four possible causes of such premature aging, calling for a more careful and thus more reliable application of protective devices.

INTRODUCTION

The great advantages of today's electronic devices in our computers and smart appliances, including their compactness, high speed and low power consumption, also make them more sensitive and vulnerable to damage or improper operation caused by disturbances such as power-line surges. These surges are caused by lightning or by switching of circuits during normal operations or during maintenance and are generally unavoidable. To protect their equipment investment and their operational continuity, users install on their power lines a variety of surge suppressors, typically based on metal-oxide varistors.

Thus, the specification of a protective varistor is driven by the wish to limit overvoltages, and a race is set up among competitors to provide the "best" protection by specifying a varistor with the lowest possible voltage rating. However, users need to know that the best approach is not necessarily the one claiming to suppress surges to the lowest level. The varistor is provided to limit surge voltages, and thus is intended to survive repeated interventions of diverting surge currents. Typical "Pulse Rating" curves published by varistor manufacturers enable varistor users to design for these repeated surge current diversions.

The greatest threat to long-term reliability of varistors might well be, not the repeated surges, but the repeated momentary overvoltages ("swells") that can occur in ac power systems[+]. There is not much information presently available on the frequency of occurrence of swells, on their duration, and on their amplitude. Anecdotal episodes are cited among varistor users as possible explanations for varistor failures, but remain undocumented. However, better characterization of swells can be expected to emerge from the expanding use of new disturbance monitors accumulating detailed information on power system anomalies. This better characterization will allow a new look at the relative contribution of swells to varistor aging. In the meantime, this paper will serve to alert the community of varistor users to the fallacy of specifying excessively low clamping voltages.

In this paper, we first review the conditions causing varistor interventions that may produce aging. Then, we report exploratory work in progress at the National Institute of Standards and Technology (NIST) on the long term effects of repeated application of swells. The motivation is twofold: to discourage indiscriminate specification of devices with excessively low voltage ratings, and to challenge manufacturers and physicists involved in the varistor technology to develop improved understanding of the phenomena leading to aging under repeated surge and swell applications.

INCREASED STRESSES FACED BY VARISTORS WITH LOWER CLAMPING VOLTAGE

1. Increased number of surge-related interventions

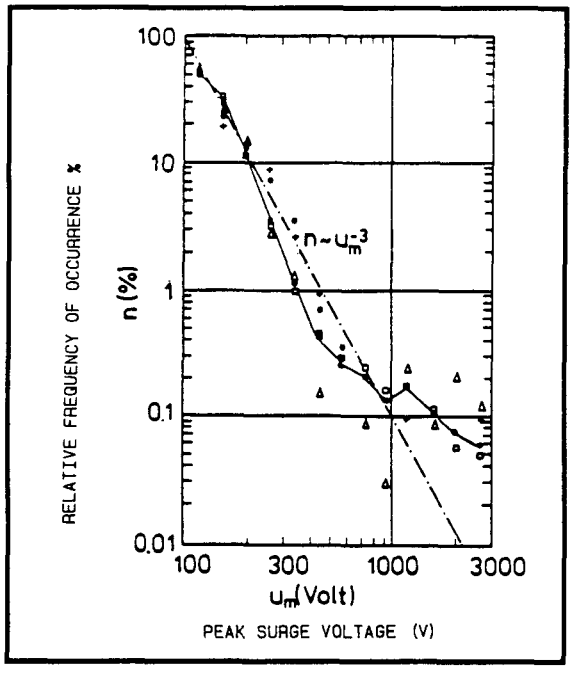
A varistor installed at an intermediate point of a power system, downstream from the service entrance, is exposed to a combination of surges: internal surges from load switching within the building, and residual surges from external lightning surges. The frequency of occurrence of these surges has been the subject of many surveys of which a recent publication is representative [2]. In all these surveys, the steep rise in the number of occurrences of surges of lower amplitude is quite remarkable. In other words, a varistor attempting to clamp at lower surge levels will be required to divert a much greater number of surges than a varistor selected with only slightly higher voltage rating. The varistor with the lower clamping voltage will expend its pulse rating faster than a varistor with higher clamping voltage.

As an example, compare the situation created for two 20-mm diameter varistors, the first one rated 250 V (the lowest possible for a 220 or 240 V circuit) the second rated 320 V (a more prudent choice according to the thesis of this paper). Consider now a surge event occurring at the 1000 V peak level as shown in the frequency of occurrence graph of Figure 1, taking this level as an *open circuit voltage* (no surge suppressor installed nearby). We can draw an equivalent circuit (Figure 2) to compute the peak surge current in a varistor installed in that environment. A source impedance of 2 Ω is taken as representative [3], so that the circuit consists of the 1000 V source and the series combination of 2 Ω and the varistor, for which the resulting current can be determined by a simple iteration, referring to the published varistor characteristics of Figure 3. For the 250 V varistor, at a peak current of about 180 A, the clamping voltage is about 650 V (180 A · 2 Ω + 650 V ≅ 1000 V). For the 320 V varistor, the peak current is only 100 A, albeit at an 800 V clamping level. Referring now to the published Pulse Rating of the

---

+ *Understanding the situation has been made more difficult because of the unfortunate use of the word "surge" in U.S. jargon by two different engineering communities, with different meanings. To the protection engineer, "surge" means a transient overvoltage with a duration of microseconds and a peak amplitude of several times the normal peak system voltage. To some electronic engineers and users of disturbance monitors, "surge" means a momentary overvoltage lasting from a few cycles to a few seconds and an amplitude ranging from a few percent above the normal peak line voltage to some level between once and twice that normal voltage. To avoid the confusion, a proposal has been made to use the word "swell" for an abbreviation of "momentary overvoltage" [1]. This paper will use "swell" to encourage the process.*

two varistors (Figure 4), for a 20 μs duration, we note a capability of 120,000 pulses at 100 A, and only 20,000 pulses at 180 A. In other words, for the same surge environment, the 250 V varistor will reach the limit of its pulse rating six times earlier than the 320 V varistor.



Source: Ref [2]

Figure 1 - Frequency of occurrence of surges in low-voltage systems

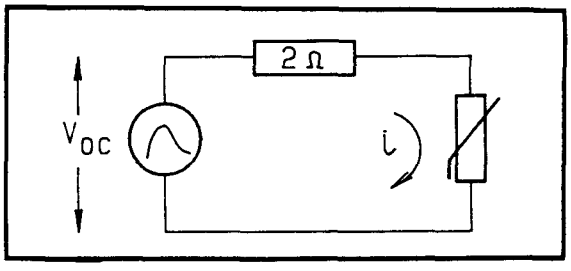
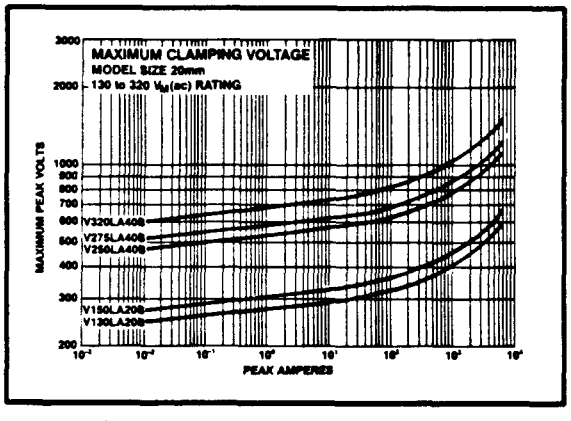
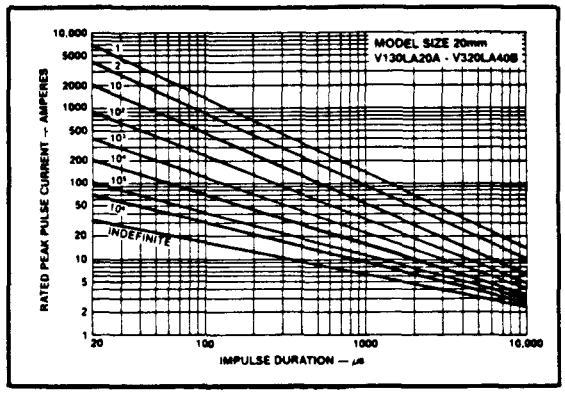


Figure 2 - Equivalent circuit for computing the surge current in a varistor



Source: Ref [10]

Figure 3 - Typical published varistor I-V characteristics



Source: Ref [10]

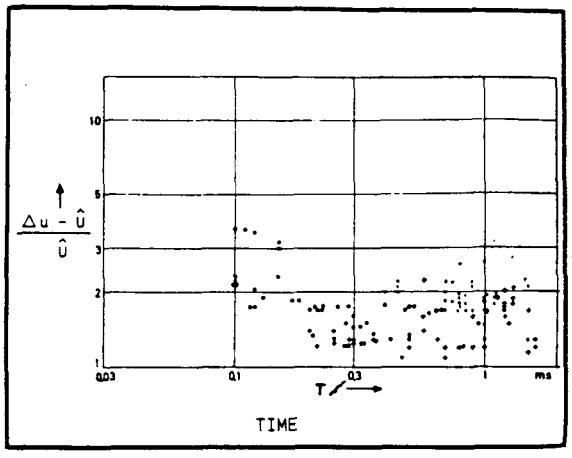
Figure 4 - Typical family of published Pulse Ratings.

2. Increased level of surge-related interventions

A varistor provided for limiting the voltage amplitude of switching surges is required to divert a substantial surge current whenever the circuit voltage exceeds the varistor clamping voltage. The energy deposited into the ceramic material increases as the surge diversion is longer when clamping occurs at a lower level of a given surge. [The fact that the clamping voltage is lower reduces the term V in the integral of $(V \cdot i \cdot dt)$, the energy expended by the varistor. however, since the lower clamping voltage is achieved by using a thinner varistor of the same diameter, the energy per volume unit of material is still increased as i is higher and the integration period is longer.] This point is most important for the longer surges encountered in the environment. For surges of shorter duration, the effect may be neglected because varistors used in an environment where more energetic surges occur are sized for the high energy surges. Reference [4] shows that a large number of low energy surges can be diverted without consuming a significant portion of the varistor pulse rating, compared to the consumption occurring for lightning surges [5] . Extreme cases of switching surge severity, such as those associated with the switching of large capacitor banks, can be a serious threat to those varistors specified for lower voltage clamping [6] .

A varistor can be exposed to surges associated with the blowing of a fuse at the end of a cable; overvoltage amplitudes can reach 2 to 2.5 times the system voltage, with durations exceeding milliseconds [7] . These surges originate from the energy stored in the cable inductance by the rising fault current until its interruption by the fuse. While this scenario is probably infrequent, it may be the most threatening to the varistor survival. Figure 5, excerpted from Ref [7], shows that the scenario of a fuse blowing at the end of a long cable can produce surges with durations on the order of milliseconds, and amplitudes reaching 2 to 3 times the system voltage peak. The current in the varistor will be determined by the rate of discharging the energy trapped in the cable inductance at the time of current interruption. This energy will be discharged into the varistor as long as the source voltage, plus the inductive voltage in the cable, exceed the varistor voltage.

To illustrate the implications, compare the situation for two 20-mm diameter varistors, one rated 250 V and the other 320 V, proposed for a 220-240 V rms system. Consider two cases of overvoltages exceeding respectively 2 times the peak system voltage (≅ 670 V) and 2.5 times (≅ 840 V) for a duration of 1 ms (these are not extreme cases according to Figure 5). Referring to Figure 3, the peak surge currents drawn by the two varistors at these two voltages are shown in the following table.



Source: Ref [7]

Figure 5 - Voltages observed at the end of a cable during fuse blowing

TABLE 1
CURRENTS ASSOCIATED WITH SURGES
OF 2 AND 2.5 TIMES THE PEAK SYSTEM VOLTAGE

	Varistor rating	
	250 V	320 V
Current at 630 V	150 A	0.2 A
Current at 780 V	1000 A	50 A

Referring to Figure 4, the pulse ratings for a 20-mm varistor at the four currents defined in the table above, with a duration of 1 ms, are the following:

1000 A: current far in excess of the rating
the varistor is destroyed.
150 A: capability of 1 pulse
the varistor survives one event only
50 A: capability of 100 pulses
the varistor outlasts fuses
0.2 A: unlimited capability

These pulse rating levels show that for a fuse blowing scenario producing 2.5 times the system overvoltage, a severe assumption, the 320 V varistor (drawing 50 A) can survive 100 such events, while the 250 V varistor (drawing 1000 A) will be destroyed. For a scenario of 2.0 times the system voltage, the 320 V varistor is in no jeopardy (unlimited capability), while the 250 V varistor can only survive one event at 150 A. For practical purposes, the 320 V varistor survives fuse blowing transients but the 250 V varistor can be expected to fail. In other words, the attempt to reduce the clamping voltage by 22% (ratio of 320 to 250) changes the life expectation in typical fuse blowing situations from adequate to assured destruction.

3. Decreased thermal runaway threshold after a surge

A varistor involved in the diversion of a surge is heated by the energy dissipated in the bulk material. The resulting increased temperature lingers after the surge, as the heat in the varistor is slowly dissipated to the environment. Because the varistor has a negative temperature coefficient of the nominal voltage in this range, the increased temperature increases the standby current in the varistor, and this standby current is much larger by the vary nature of the nonlinear behavior of the varistor for a varistor of lower voltage rating. The temperature of the varistor after the surge reflects the decay of the transient temperature reached at the end of the surge current pulse, but augmented by the increasing steady-state heating resulting from the higher standby current. A point-of-no-return can be reached where the increase exceeds the decay, resulting in a thermal runaway and destruction of the varistor [8]. For a varistor with a low voltage rating, this point of thermal runaway is reached for a lower surge current - its capability for recovering from a unusually large surge is much less than that of a varistor with only slightly higher voltage rating. Just short of the thermal runaway, accelerated aging can be expected, even if the varistor recovers from the close encounter with disaster.

4. Increased stresses during swells

A varistor has the difficult task of suppressing brief surges, but NEVER attempting to suppress the longer swells that inescapably occur in the course of operating a large power system. If, in a misguided attempt to "better" suppress surges, the suppression level is chosen too low, the surge protector will attempt to suppress the power system swells and be destroyed in the process. For swells that would not produce destruction during a one-time occurrence, long term effects of repeated swells might produce premature aging. Typical sources of short-duration swells are power system recovery from disturbances and load shedding. Long-duration swells are associated with unbalanced phase voltages occurring upon loss of the neutral conductor in three-phase systems or in center-tapped, single-phase systems. Malfunction of voltage regulators is another cause of long-duration swells.

SHORT AND LONG TERM EFFECTS OF SWELLS

While the effect of repeated surge diversion is documented by the Pulse Rating curves published by manufacturers, the short and long term effects of swells are not documented. Only the line-voltage ratings assigned by the varistor manufacturer to a given piece of ceramic material imply sufficient margins to provide long-term reliability of the varistors exposed to swells. Considering the uncertainty of actual limits in the number, duration, and amplitude of swells (as opposed to expectations raised by some standards [9]) on the one hand, and the competitive drive toward lower clamping voltages on the other hand, a closer examination of the issue is in order.

To that effect, exploratory work is in progress at NIST, including experiments and computer modeling. The experiments involve exposing varistors to well-controlled swells created by a sinusoidal voltage source with low internal impedance. The amplitude and the duration of the swells can be programmed to apply swells at predetermined percentages of the nominal voltage of each varistor specimen. The model is based on an equivalent circuit that postulates no change in the varistor I-V characteristic during the occurrence of a single swell. This model makes possible a prediction of the power dissipation and temperature rise in the varistor for any arbitrary swell amplitude and duration. One of the long range objectives of the investigation is to detect changes in the V-I characteristics after repeated application of swells, leading to a family of "Swell Ratings" similar to the published Pulse Ratings.

Modeling of varistor during swells

An electrical circuit analysis model of a metal-oxide varistor was used to investigate the behavior of varistors when exposed to swells [10] . The model consists of two basic parts: The electrical part of the model predicts the current-voltage (I-V) behavior of the varistor and the thermal part of the model predicts the temperature changes caused by the energy deposited in the varistor during the swell.

Figure 6 shows the electrical and thermal circuits of the model. The equivalent electrical circuit of the varistor, shown on the left-hand side of the figure, is composed of five electrical elements along with a voltage generator, Vg. The major elements comprising the model of the varistor are the nonlinear resistance (S) and the bulk resistance of the varistor material (Ron). The current through the varistor is modeled as a nonlinear voltage-dependent resistance which obeys the classical varistor equation $Imov = k \cdot (Vmov)^{\alpha}$. The electrical and thermal simulations were performed using the SABER [11] software system , which uses behavioral equations to mimic nearly any physical device or process.

In the context of this paper, aimed at alerting the engineering community to the risks of excessively low clamping voltage, the model computations were performed for a varistor rated 130 V rms. A varistor rated 130 V rms is the lowest rating offered by manufacturers for a 120 V circuit. To represent a possible worst case, the varistor characteristics used for the computations were set at the lower limit of the tolerance band. Similar conclusions would be obtained for a varistor rated 250 V, the lowest rating offered for 220 V circuits. The values of k and α in the varistor equation were obtained by fitting the equation to published maximum I-V points for 1 mA and 100 A; then, keeping the same value of α, a new value of k was computed for the lower-limit 130 V rms varistor.

The remaining elements included in the electrical model are the inductor (L), associated with the inductance of the lead wires (estimated to be 0.01 nH); the resistor (Roff) that accounts for the standby current of the varistor in the low current density region (estimated to be 1000 MΩ); and the capacitor (C), which represents the inherent capacitance of the varistor element (estimated to be 0.002 μF). For modeling at dc or power-line frequencies, neither the inductor nor the capacitor play a significant role, but they were included in the model to allow transient response modeling when desired.

The thermal model of the varistor is depicted on the right side of Figure 6. This portion of the figure shows a varistor, 22 mm in diameter, with radial copper leads, 0.81 mm in diameter, attached to a thermal sink 10 mm from the varistor body. The thermal capacity of the varistor is calculated for a varistor weight of 3 g and composed entirely of zinc oxide. Thus, the thermal circuit consists of the thermal resistances, θl, of the two copper leads in parallel with the thermal capacitance Ct. To simplify the model, no allowance is made for heat loss by convection, the situation that can be expected in a tight packaging of a varistor without any encapsulation. Heat losses by radiation are negligible at the temperatures involved. These postulates provide worst case results, a justifiable approach when considering reliability. The thermal generator consists of the power dissipated in the electrical components enclosed by the dotted lines. The total power dissipated in the varistor, P, is calculated as the sum of the power dissipated in the nonlinear resistance and the power dissipated in the resistor Ron and may be written as $P = (Imov \cdot Vmov) + (Imov^2 \cdot Ron)$. From this model, both the electrical and the thermal behavior of a varistor may be obtained.

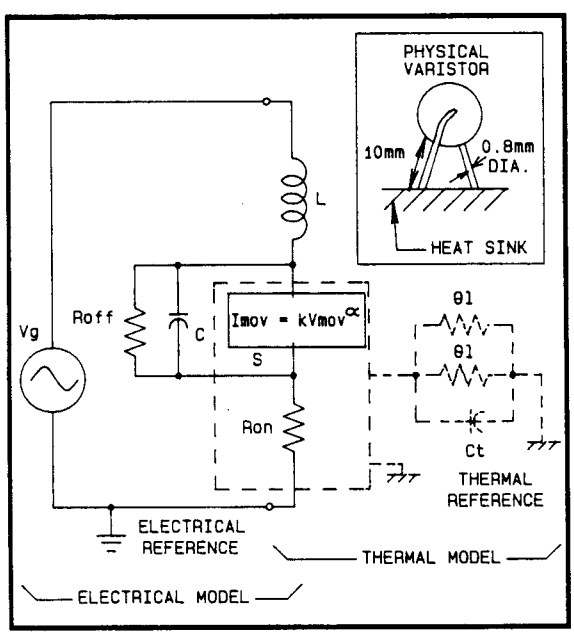


Figure 6 - Electrical and thermal model of a varistor

Figure 7 shows the dc I-V characteristic of the varistor, as predicted by the model, over a range of 1 μA to 1000 A. The 1 mA and 100 A points, of course, match the published characteristic from which k and α were derived. The upturn, starting at about 80 A, reflects the additional voltage contributed by the series resistance (Ron). This plot illustrates both the basic desired characteristic of a varistor and the undesirable implication for swell conditions. The desired characteristic, of course, is the small change of terminal voltage over a wide range of surge currents, which results in the clamping effect. The undesirable implication is that a varistor will exhibit large increases in standby current, with corresponding power dissipation and temperature rise, for relatively small increases in applied voltage, the swell condition.

Figure 8 shows the current and the temperature rise predicted by the model for three voltage levels applied to a varistor with a rating of 130 V rms at the lower limit of tolerance. The waveforms were simulated for the first 35 ms (approximately two power-line cycles). In Figure 8(A), the applied voltage is 120 V rms and the current through the varistor (solid line) leads the voltage (dotted line) since most of the current is through the shunt capacitance. The small peaks on the current waveform occur at the peak values of the applied voltage and represent the contribution of the current from the nonlinear varistor element. During the time when the nonlinear varistor element passes significant current, power is dissipated in the varistor and an increase in temperature (dashed line) is observed. Figures 8(B) and 8(C) are plots of the behavior of the same varistor at 132 V rms and 145 V rms, respectively. As power is dissipated, cycle after cycle during a swell, the temperature of the varistor increases, until an equilibrium is reached between the energy input (power dissipation) and the energy output (heat losses by conduction).

Figure 9 shows the temperature rise predicted for the application of swells at two voltage levels, each with a duration of 200 s. The 200 s duration was selected because it approximates the time required to reach thermal equilibrium in this model, not because of an inference that power system swells last 200 s. The lower curve shows the temperature rise of approximately 130 °C expected from a 156 V rms swell applied to a 130 V varistor. The upper curve shows the

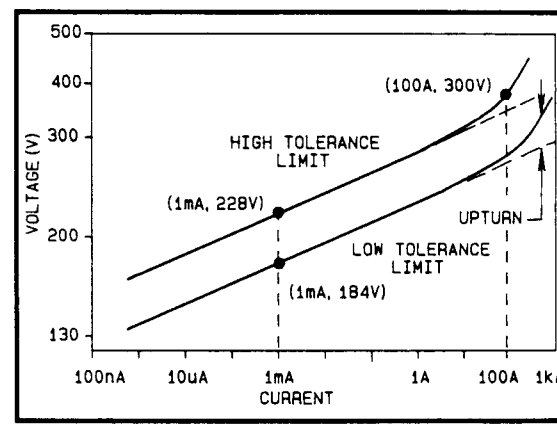


Figure 7 - I-V characteristic for a 130 V rms varistor, upper and lower tolerance limits.

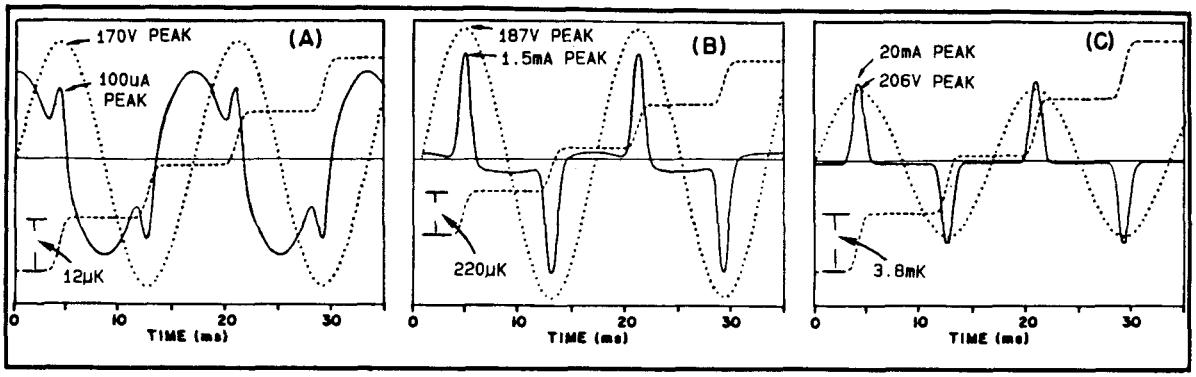


Figure 8 - Varistor current and temperature increase rate at normal line voltage (A) and during the initial portions of a 110% swell (B) and a 121% swell (C)

temperature rise of nearly 300 °C for a 160 V rms swell. Such a swell would destroy a "real" varistor. The curves illustrate that a relatively small change in the swell level, from 156 V to 160 V, produces a drastic increase in the power dissipation and the temperature rise of a varistor. Note that typical varistor specifications show a power derating starting at 85 °C ambient, and zero power rating at 125 °C. Thus, the capability of a varistor to remain within its ratings is eroded when heated by swells. This effect has to be included in the evaluation of recovery from a surge mentioned above as the third stress faced by varistors with low clamping voltages.

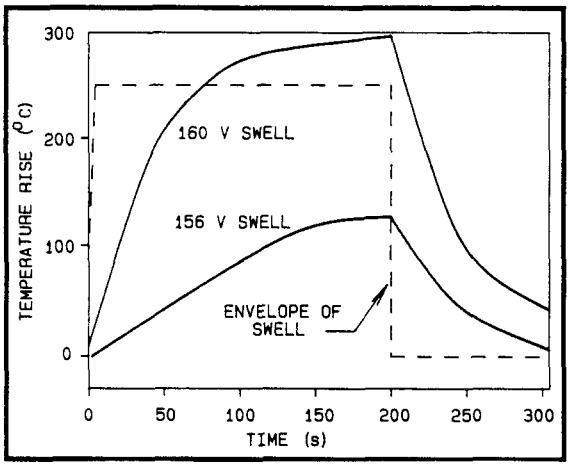


Figure 9 - Temperature rise of the internal varistor during swells

These results are applicable to a model varistor where k and α are unaffected by temperature changes. Published information on temperature coefficient states a negative voltage coefficient of -0.05% per °C at 1 mA (a better specification would be to prescribe current density rather than absolute current), with increasing coefficient below 1 mA and negligible coefficient above 1 mA. A further step in the sophistication of the model would be to adjust the parameters k or α, or both, as the temperature rises. However the concern here is primarily for swells resulting in currents above 1 mA so that the model is adequate for this purpose.

Experiments on swell applications

Several factors are known or suspected in the behavior of actual varistors, complicating the apparent simplicity of varistor modeling. Aging -- a change in the varistor characteristics in response to overstresses -- is only one of them. Even when considering only aging, there is no universally recognized criterion of degradation. Most varistor specifications refer to a ±10% limit in changes of the nominal voltage of the varistor. However, since this nominal voltage is defined arbitrarily for a 1 mA current, regardless of the varistor cross section, the criterion varies with varistor sizes. Another proposed criterion, power dissipation at some ac voltage, is complicated by the fact that the power dissipation changes with the duration of application of the ac voltage. This effect which permanently changes the characteristics of the varistor, sometimes called "formation," was quite apparent when successive swells were applied. During application of one swell, for instance, the increasing temperature of the varistor is reflected in the increased peak amplitude of the current, as shown in Figure 10. However, when comparing the pattern of increasing current peaks within a swell to the pattern of a later swell, after many additional swells, the most obvious change is a reduction of the current peaks between the original swell and the later swell. Thus, a systematic method to account for that "formation" effect must be developed to define a valid criterion of aging before conclusions can be drawn on quantifying the aging effect of swells.

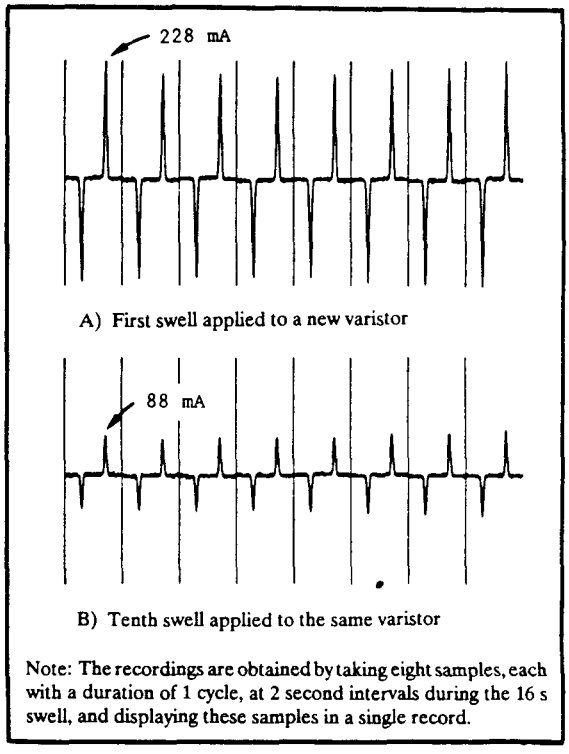


Figure 10 - Samples of varistor current during a 16 s swell at 117% of the varistor nominal voltage (198 V peak).

CONCLUSIONS

Several mechanisms involving surges or momentary overvoltages can cause accelerated, or premature aging of varistors, if the clamping voltage is selected at too low a level without appropriate consideration of all factors.

The first aging mechanism, repeated surge diversion interventions, has been well documented by the manufacturers. A low clamping level will invite more frequent interventions, but information is readily available on this mechanism. Careful designers can use the information to ensure reliability for specific environments and desired useful life.

A second mechanism, fortunately not occurring too frequently, involves fuse blowing and can produce immediate destruction of the varistor at the first occurrence if the clamping level is selected at too low a level. The implications of this situation needs greater recognition among varistor users.

A third mechanism, decreased thresholds of thermal runaway in the long term, is directly related to the selected clamping level, with aging accelerated by a low clamping level selection. This situation is well recognized by high-voltage arrester designers, but not by low-voltage electronic circuit designers.

A fourth mechanism, repeated conduction of currents associated with momentary system overvoltages ("swells"), has not been documented but is now being investigated. The results of exploratory investigations will be published when completed, to act as a catalyst for further investigations at NIST as well as by other varistor users.

The obvious, but difficult remedy to this situation is to design equipment with a reasonably high surge withstand capability so that retrofit using protective devices with very low clamping voltage will not be necessary. For those situations where a close protection would be required, a very careful consideration of all factors becomes imperative, rather than cookbook application of protective devices.

In the absence of a demanding retrofit challenge, there is no advantage and a considerable penalty in providing a too narrow protection margin by specifying needlessly low clamping voltages. Such low voltages are counterproductive to total system reliability. Thoughtful design can provide good performance with good reliability; short-term perspective and quick fixes can only compromise long-term reliability.

ACKNOWLEDGMENTS

The exploratory investigations reported in this paper were performed by a team consisting of the two authors, supported by the experimental contributions of Mark Parker and Albert Cooper. Funding of the work by the National Institute of Standards and Technology was augmented by an informal consortium of private sector sponsors, organized by the Building Industry Consulting Service International.

REFERENCES

[1] Martzloff, F.D., and Gruzs, T.M., "Power Quality Site Surveys: Facts, Fiction, and Fallacies," IEEE Transactions, Vol. IA-24, No. 6 November/December 1988.

[2] Goedbloed, J.J. "Transients in Low-Voltage Supply Networks," IEEE Transactions , Vol. EMC-29, No.2, May 1987.

[3] IEC Draft Document 28A(Secretariat)30, "Explanation of Interfaces for Overvoltage Categories," 1987.

[4]Korn, S., "Transients and Transient Suppressors for Power Conversion Equipment," INTELEC Conference Record, 1982.

[5] Martzloff, F.D., "The Effect of Repetitive Surges on Transient Overcurrent and Overvoltage Protective Devices," Proceedings, 5th Symposium on EMC, 1983.

[6] Martzloff, F.D., "Varistor versus Environment: Winning the Rematch," IEEE Transactions, Vol. PWRD-1, No. 2, April 1986.

[7] Meissen, W., "Überspannungen in Niederspannungsnetzen," Elektrotechnische Zeitschrift, Vol. 104, 1983.

[8] Sakshaug, E.C., Kresge, J.S., and Miske,. S.A., "A New Concept in Station Arrester Design," IEEE Transactions, Vol. PAS-96, No. 2, March/April 1977.

[9] ANSI C84.1-1982, American National Standard Voltage Ratings for Electric Power Systems and Equipment.

[10] General Electric Company, Transient Voltage Suppression, Fifth Edition, 1986, Liverpool, NY, p. 3-39.

[11] SABER is a product of Analogy, Inc., Beaverton, OR.