Introduction to Social Macrodynamics. Compact Macromodels of the World System Growth by Andrey Korotayev , Artemy Malkov, and Daria Khaltourina. Moscow: Editorial URSS, 2006. Pp. 21–33.

Chapter 2

A Compact Macromodel of World Population Growth

The fact that up to the 1960s world population growth had been characterized by a hyperbolic trend was discovered quite long ago (see, *e.g.*, von Foerster, Mora, and Amiot 1960; von Hoerner 1975; Kremer 1993; Kapitza 1992, 1999, *etc.*). In 1960 von Foerster, Mora, and Amiot conducted a statistical analysis of the available world population data and found out that the general shape of the world population (N) growth is best approximated by the curve described by the following equation:

$$N = \frac{C}{t_0 - t}, \qquad (2.1)$$

where C and t_0 are constants, whereas t_0 corresponds to an absolute limit of such a trend at which N would become infinite, and thus logically implies the certainty of the empirical conclusion that further increases in the growth trend will cease well before that date, which von Foerster wryly called the "doomsday" implication of power-law growth (he refers tongue-in-cheek to the estimated t_0 as "Doomsday, Friday, 13 November, A.D. 2026").

Von Foerster, Mora, and Amiot try to account for their empirical observations by modifying the usual starting equations (0.1) and (0.3) for population dynamics, so as to describe the process under consideration:

$$\frac{dN}{dt} = B - D, \qquad (0.1)$$

where N is the number of people, B is the number of births, and D is the number of deaths in the unit of time;

$$\frac{dN}{dt} = (a_1 N) - (a_2 N + bN^2), \tag{0.3}$$

where a_1N corresponds to the number of births *B*, and $a_2N + bN^2$ corresponds to the number of deaths in equation (0.1); let us recollect that *r*, *K*, a_1, a_2, *b* are positive coefficients connected between themselves by the following relationships:

$$r = a_1 - a_2 \quad \text{and} \quad b = \frac{r}{K}, \tag{0.4}$$

They start with the observation that when individuals in a population compete in a limited environment, the growth rate typically *decreases* with the greater number *N* in competition. This situation would typically apply where sufficient communication is lacking to enable resort to other than a competitive and nearly zero-sum multiperson game. It might not apply, they suppose, when the elements in a population "possess a system of communication which enables them to form coalitions" and especially when "all elements are so strongly linked that the population as a whole can be considered from a game-theoretical point of view as a single person playing a two-person game with nature as the opponent" (von Foerster, Mora, and Amiot 1960: 1292). Thus, the larger the population (N^k coalition members, where $k \leq 1$) the more the decrease of natural risks and the higher the population growth rate. They suggest modeling such a situation through the introduction of nonlinearity in the following form:

$$\frac{dN}{dt} = (a_0 N^{\frac{1}{k}}) N, \tag{2.2}$$

where a_0 and *k* are constants, which should be determined experimentally. The analysis of experimental data by von Foerster, Mora, and Amiot determines values $a_0 = 5.5 \times 10^{-12}$ and $k = 0.99$ that produce the hyperbolic equation for world population growth:

$$N = N_1 \left(\frac{t_0 - t_1}{t_0 - t} \right)^k, \tag{2.3}$$

which, assuming $k = 1.0$ (von Hoerner1975) is written more succinctly as (2.1) and in equivalent form (Kapitza 1992, 1999) as (2.4):[1]

$$\frac{dN}{dt} = \frac{N^2}{C}. \tag{2.4}$$

Though von Foerster's, von Hoerner's and Kapitza's models produce a phenomenal fit with the empirical data, they do not account for mechanisms of the hyperbolic trend; as we shall see in the next chapter, Kremer's (1993) model accounts for it, but it is rather complex. In fact, the general shape of world population growth dynamics could be accounted for with strikingly simple models like the one we would like to propose ourselves below (or the model proposed by Tsirel [2004]).[2]

With Kremer (1993), Komlos, Nefedov (2002) and others (Habakkuk 1953; Postan 1950, 1972; Braudel 1973; Abel 1974, 1980; Cameron 1989; Artzrouni and Komlos 1985 *etc.*), we make "the Malthusian (1978) assumption that population is limited by the available technology, so that the growth rate of population is proportional to the growth rate of technology" (Kremer 1993: 681–2),[3] and that, on the other hand, "high population spurs technological change because it increases the number of potential inventors…[4] In a larger population there will be proportionally more people lucky or smart enough to come up with new ideas"[5] (Kremer 1993: 685), thus, "the growth rate of technology is proportional to total population"[6] (Kremer 1993: 682; see also, *e.g.*, Kuznets 1960; Grossman and Helpman 1991; Aghion and Howitt 1992, 1998; Simon 1977, 1981, 2000; Komlos and Nefedov 2002; Jones 1995, 2003, 2005 *etc.*).

---

[1] See Appendix 3 for more detail.

[2] For other models of the world population hyperbolic growth see Cohen 1995; Johansen and Sornette 2001; Podlazov 2004.

[3] In addition to this, the absolute growth rate is proportional to population itself – with the given relative growth rate a larger population will increase more in absolute numbers than a smaller one.

[4] "This implication flows naturally from the nonrivalry of technology… The cost of inventing a new technology is independent of the number of people who use it. Thus, holding constant the share of resources devoted to research, an increase in population leads to an increase in technological change" (Kremer 1993: 681).

[5] The second assumption is in fact Boserupian rather than Malthusian (Boserup 1965; Lee 1986).

[6] Note that "the growth rate of technology" means here the relative growth rate (*i.e.*, the level to which technology will grow in the given unit of time in proportion to the level observed at the beginning of this period). This, of course, implies that the absolute speed of technology growth in the given period of time will be proportional not only to the population size, but also to the absolute technology level at the beginning of this period.

The simplest way to model mathematically the relationships between these two subsystems (which, up to our knowledge, has not yet been proposed)[7] is to use the following set of differential equations:

$$\frac{dN}{dt} = a(bK - N)N \;, \tag{2.5}$$

$$\frac{dK}{dt} = cNK \;, \tag{2.6}$$

where N is the world population, K is the level of technology; bK corresponds to the number of people (N), which the earth can support with the given level of technology (K). With such a compact model we are able to reproduce rather well the long-run hyperbolic growth of world population before 1962-3.

With our two-equation model we start the simulation in the year 1650 and do annual iterations with difference equations derived from the differential ones:

$K_{i+1} = K_i + cN_iK_i$,
$N_{i+1} = N_i + a(bK_{i+1} - N_i)N_i$.

We choose the following values for the constants and initial conditions: $N = 0.0545$ of tens of billions (*i.e.* 545 million)[8]; $a = 1$; $b = 1$; $K = 0.0545$;[9] $c = 0.05135$. The outcome of the simulation, presented in Diagrams 2.1–2 indicates that irrespective of its simplicity the model is actually capable of replicating quite reasonably the population estimates of Kremer (1993), US Bureau of the Census (2004) and other sources (Thomlinson 1975; Durand 1977; McEvedy and Jones 1978: 342–51; Biraben 1980; Haub 1995: 5; UN Population Division 2005; World Bank 2005) in most of their characteristics and in terms of the important turning points:

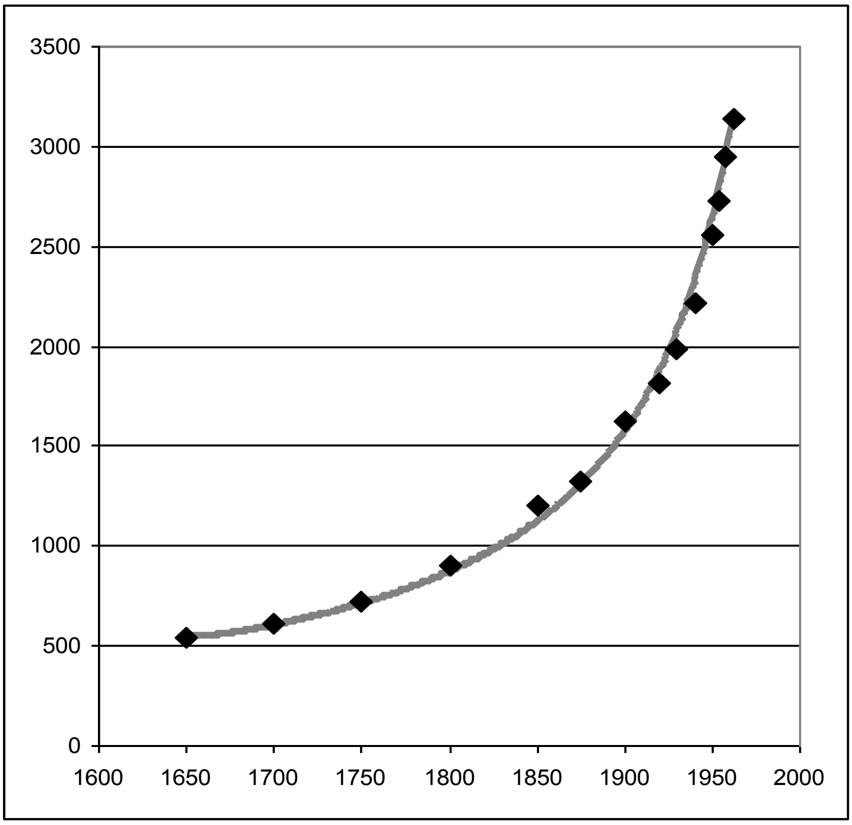
---

[7] The closest proposed model is the one by Tsirel (2004); see our discussion of this very interesting model in Korotayev, Malkov, and Khaltourina 2005: 38–57.

[8] We chose to calculate the world population in tens of billions (rather than, say, in millions) to minimize the rounding error stemming from discrete computer nature (which was to be taken most seriously into account in our case, as the object of modeling had evident characteristics of a blow-up regime).

[9] To simplify the calculations we chose value "1" for both a and b; thus, K in our simulations was measured directly as the number of people which can be supported by the Earth with the given level of technology.

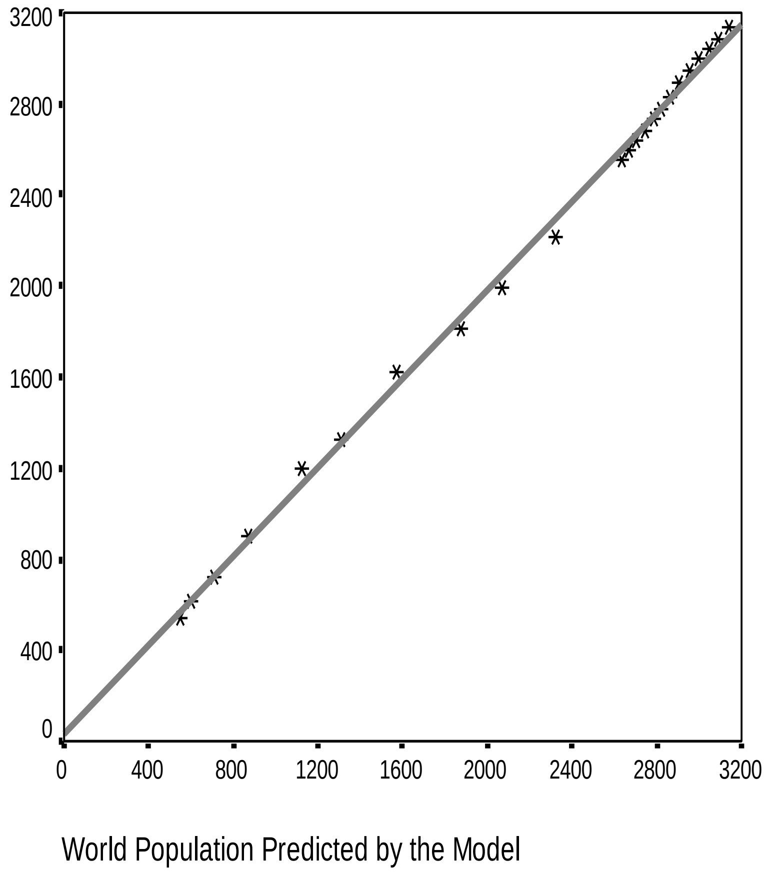
Diagram 2.1. Predicted and Observed Dynamics of the World Population Growth, in millions (1650–1962 CE)



NOTE: The solid grey curve has been generated by the model; black markers correspond to the estimates of world population by Kremer (1993) for pre-1950 period, and US Bureau of Census (2005) world population data for 1950–1962.

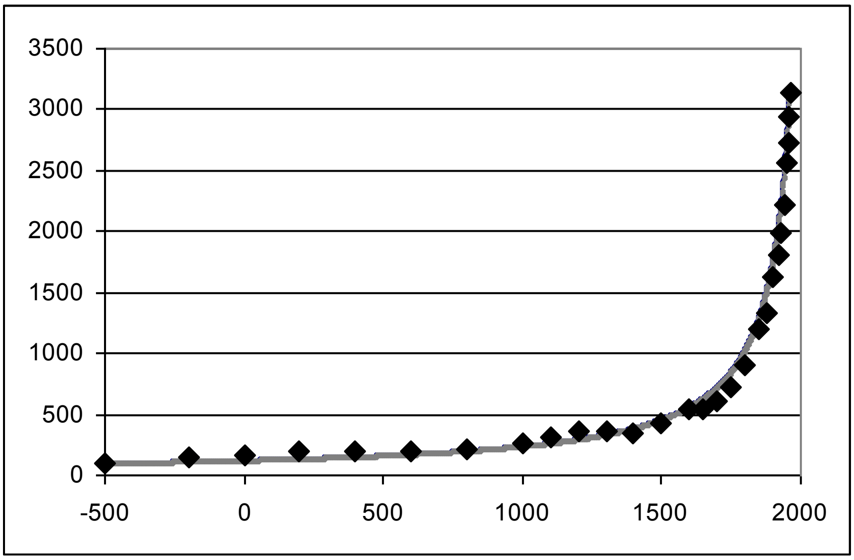
The correlation between the predicted and observed values for this simulation looks as follows: $R = 0.9989$, $R^2 = 0.9978$, $p << 0.0001$, which, of course, indicate an unusually high fit for such a simple model designed to account for demographic macrodynamics of the most complex social system (see Diagram 2.2):

Diagram 2.2. Correlation between Predicted and Observed Values (1650–1962)



We start our second simulation in the year 500 BCE. In this case we choose the following values of the constants and initial conditions: N = 0.01 of tens of billions (*i.e.* 100 million); a = 1*;* b = 1*;* K = 0.01; c = 0.04093. The outcome of the simulation, presented in Diagrams 2.3–4 indicates that irrespective of its extreme simplicity the model is still quite capable of replicating rather reasonably the population estimates of Kremer (1993), US Bureau of the Census (2004) and other sources in most of their characteristics and in terms of the important turning points even for such a long period of time:

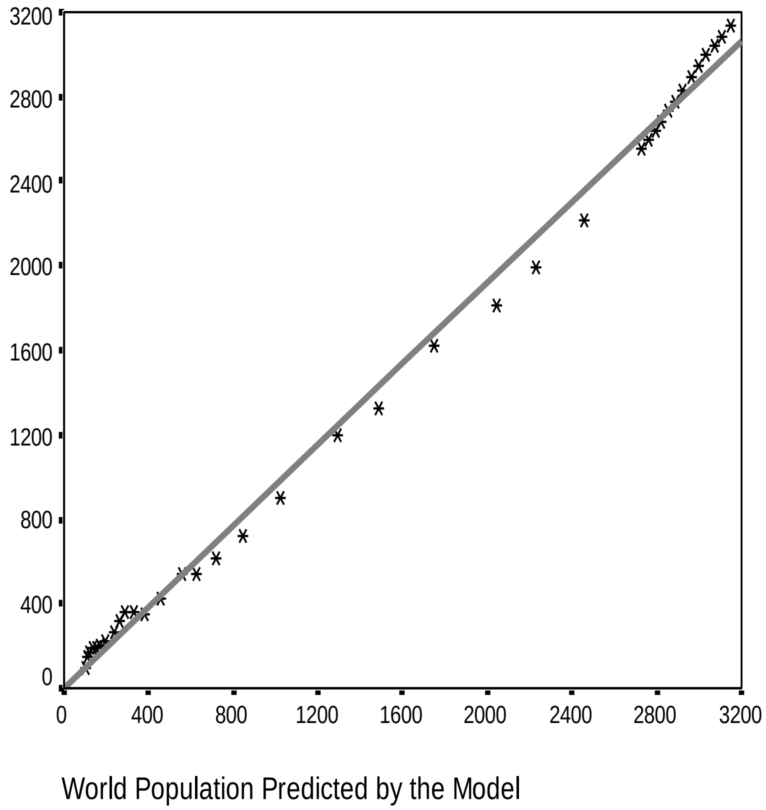
Diagram 2.3. Predicted and Observed Dynamics of the World Population Growth, in millions (500 BCE – 1962 CE)



NOTE: The solid grey curve has been generated by the model; black markers correspond to the estimates of world population by Kremer (1993) for pre-1950 period, and US Bureau of Census (2005) world population data for 1950–1962.

The correlation between the predicted and observed values for this simulation looks as follows: R = 0.9983, R^2 = 0.9966, p << 0.0001, which, of course, again indicate an unusually high fit for such a simple model designed to account for demographic macrodynamics of the most complex social system for *c*. 2500 years (see Diagram 2.4):

Diagram 2.4. Correlation between Predicted and Observed Values



Note that even when the simulation was started *c*. 25000 BCE, it still produced a fit with observed data as high as 0.981 (R^2 = 0.962, p << 0.0001).[10]

Thus, it turns out that the set of two differential equations specified above accounts for 96.2 per cent of all the variation in demographic macrodynamics of the world in the last 25 millennia; it also accounts for 99.66% of this macrovariation in 500 BCE – 1962 CE, and it does for 99.78% in 1650–1962 CE.

---

[10] The simulation was started in the 24939 BCE and done with 269 centennial iterations ending in 1962 CE. In this case we chose the following values of the constants and initial conditions: N = 0.00334 billion (*i.e.* 3.34 million); a = 1; b = 1; K = 0.00334; c = 2.13.

In fact, we believe this may not be a coincidence that the compact macromodel shows such a high correlation between the predicted and observed data just for 500 BCE – 1962 CE. But why does the correlation significantly decline if the pre-500 BCE period is taken into account?

To start with, when we first encountered models of world population growth, we felt a strong suspicion about them. Indeed, such models imply that the world population can be treated as a system. However, at a certain level of analysis one may doubt if this makes any sense at all. The fact is that up until recently (especially before 1492) humankind did not constitute any real system, as, for example. The growth of the Old World, New World, Australia, Tasmania, or Hawaii populations took place almost perfectly independently of each other. For example, it seems entirely clear that demographic processes in, say, West Eurasia in the 1st millennium CE did not have the slightest impact on the demographic dynamics of the Tasmanian population during the same time period.

However, we believe that the patterns observed in pre-Modern world population growth are not coincidental at all. In fact, they reflect population dynamics of quite a real entity, the World System. We are inclined to speak, with Andre Gunder Frank (*e.g.*, Frank and Gills 1994) but not with Wallerstein (1974), about a single World System which originated long before the "long 16th century".

Note that the presence of a more or less well integrated World System, comprising most of the world population, is a necessary pre-condition for the high correlation between the world population numbers generated by our model and the observed ones. For example, suppose we encounter a case when the world population of *N* grew 4-fold but got split into 4 perfectly isolated regional populations comprising *N* persons each. Of course, our model predicts that a 4-fold increase of the world population would tend to lead to a 4-fold increase in the relative world technological growth rate. But have we any grounds to expect to find this in the case specified above? Of course not. Yes, even in this case a four times higher number of people are likely to produce 4 times more innovations. However, the effect predicted by our model would be only observed if innovations produced by any of the four regional populations were shared among all the other populations. However, if we assumed that the four respective populations lived in perfect isolation from each other, then such sharing would not take place, and the expected increase in technological growth rate would not be observed, thereby producing a huge gap between the predictions generated by our model and actually observed data.

It seems that this was just the 1st millennium BCE when the World System integration reached a qualitatively new level. A strong symptom of this seems to be the "Iron Revolution", as a result of which the iron metallurgy spread within a few centuries (not millennia!) throughout a huge space stretching from the Atlantic to the Pacific, producing (as was already supposed by Jaspers

[1953]) a number of important unidirectional transformations in all the main centers of the emerging World System (the Circummediterranean region, Middle East, South Asia, and East Asia), after which the development of each of those centers cannot be adequately understood, described and modeled without taking into consideration the fact that it was a part of a larger and perfectly real whole – the World System.

A few other points seem to be relevant here. Of course, there would be no grounds to speak about the World System stretching from the Atlantic to the Pacific even at the beginning of the 1st Millennium CE if we applied the "bulk-good" criterion suggested by Wallerstein (1974), as there was no movement of bulk goods at all between, say, China and Europe at this time (as we have no grounds not to agree with Wallerstein in his classification of the 1st century Chinese silk reaching Europe as a luxury, rather than a bulk good). However, the 1st century CE (and even the 1st millennium BCE) World System would definitely qualify as such if we apply a "softer" information network criterion suggested by Chase-Dunn and Hall (1997). Note that at our level of analysis, the presence of an information network covering the whole World System is a perfectly sufficient condition, which makes it possible to consider this system as a single evolving entity. Yes, in the 1st millennium BCE any bulk goods could hardly penetrate from the Pacific coast of Eurasia to its Atlantic coast. However, by that time the World System had reached such a level of integration that, say, iron metallurgy could spread through the whole World System within a few centuries.

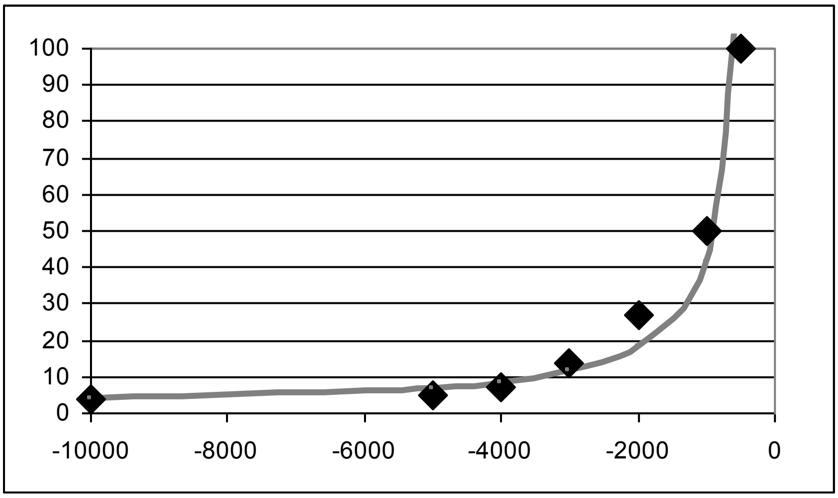
The other point is that even in the 1st century CE the World System still covered far less than 50% of all the Earth's terrain. However, what seems to be far more important is that already by the beginning of the 1st century CE more than 90% of all the world population lived in just those regions which were constituent parts of the 1st century CE World System (the Circummediterranean region, Middle East, South, Central and East Asia) (see, *e.g.*, Durand 1977: 256). Hence, since the 1st millennium BCE the dynamics of world population reflects very closely just the dynamics of the World System population.

On the other hand, it might not be coincidental that the hyperbolic growth trend may still be traced back to 25000 BCE. Of course, we do not insist on the existence of anything like the World System, say, around 15000 BP. Note, however, that there does not seem to be any evidence for hyperbolic world population growth in 40000 – 10000 BCE. In fact the hyperbolic effect within the 25 millennia BCE is produced by world population dynamics in the last 10 millennia of this period that fits the mathematical model specified above rather well (though not as well, as the world population dynamics in 500 BCE – 1962 CE [let alone 1650 – 1962 CE]).

The simulation for 10000 – 500 BCE was done with the following constants and initial conditions: N = 0.0004 of tens of billions (*i.e.* 4 million); a = 1; b = 1; K = 0.0004; c = 0.32.

The outcome of the simulation, presented in Diagram 2.5 indicates that the model is still quite capable of replicating rather reasonably the population estimates of McEvedy and Jones (1978) and Kremer (1993) for the 10000 – 500 BCE period:

Diagram 2.5. Predicted and Observed Dynamics of the World Population Growth, in millions (10000 – 500 BCE)



NOTE: The solid grey curve has been generated by the model; black markers correspond to the estimates of world population by McEvedy and Jones (1978) and Kremer (1993).

The correlation between the predicted and observed values for this simulation looks as follows: $R = 0.982$, $R^2 = 0.964$, $p = 0.0001$. Note that though this correlation for 10000 – 500 BCE remains rather high, it is substantially weaker[11] than the one observed above for the 500 BCE – 1962 CE and, especially, 1650–1962 CE (in fact this is visible quite clearly even without special statistical analysis in Diagrams 2.1, 2.3, and 2.5). On the one hand, this result could hardly be regarded as surprising, because it appears evident that in 10000 – 500 BCE the World System was much less tightly integrated than in 500 BCE – 1962 CE (let alone in 1650–1962 CE). What seems more remarkable is that for 10000 – 500 BCE the best fit is achieved with a substantially different value of the coefficient *c*, which appears to indicate that the World System devel-

[11] Note, however, that even for 10000 – 500 BCE our hyperbolic growth model still demonstrates a much higher fit with the observed data than, for example, the best-fit exponential model ($R^2 = 0.737$, $p = 0.0003$).

opment pattern in the pre-500 BCE epoch was substantially different from the one observed in the 500 BCE – 1962 CE era, and thus implies a radical transformation of the World System in the 1st millennium BCE..

We believe that among other things the compact macromodel analysis seems to suggest a rather novel approach to World System analysis. The hyperbolic trend observed for world population growth after 10000 BCE mostly appears to be a product of the growth of the World System, which seems to have originated in West Asia around that time in direct connection with the Neolithic Revolution. The presence of the hyperbolic trend indicates that the major part of the entity in question had some systemic unity, and, we believe we have evidence for this unity. Indeed, we have evidence for the systematic spread of major innovations (domesticated cereals, cattle, sheep, goats, horses, plow, wheel, copper, bronze, and later iron technology, and so on) throughout the whole North African – Eurasian Oikumene for a few millennia BCE (see, *e.g.*, Chubarov 1991; Diamond 1999 *etc.*). As a result, already at this time the evolution of societies in this part of the world cannot be regarded as truly independent. By the end of the 1st millennium BCE we observe a belt of cultures stretching from the Atlantic to the Pacific with an astonishingly similar level of cultural complexity based on agriculture involving production of wheat and other specific cereals, cattle, sheep, goats, plow, iron metallurgy, professional armies with rather similar weapons, cavalries, developed bureaucracies and so on – this list can be extended for pages. A few millennia before we would find a belt of societies with a similarly strikingly close level and character of cultural complexity stretching from the Balkans to the Indus Valley borders (note that in both cases the respective entities included the major part of the contemporary world population). We would interpret this as tangible results of the World System functioning. The alternative explanations would involve a sort of miraculous scenario – that cultures with strikingly similar levels and character of complexity somehow developed independently from each other in a very large but continuous zone, whereas nothing like them appeared in other parts of the world, which were not parts of the World System. We find such an alternative explanation highly implausible.

It could be suggested that within a new approach the main emphasis would be moved to the generation and diffusion of innovations. If a society borrows systematically important technological innovations, its evolution already cannot be considered as really independent, but should rather be considered as a part of a larger evolving entity, within which such innovations are systematically produced and diffused. The main idea of the world-system approach was to find the evolving unit. The basic idea was that it is impossible to account for the evolution of a single society without taking into consideration that it was a part of a larger whole. However, traditional world-system analysis concentrated on bulk-good movements, and core – periphery exploitation, thoroughly neglecting the above-mentioned dimension. However, the information network turns out to be the oldest mechanism of the World System integration, and remained extremely important throughout its whole history, remaining important up to the present. It seems to be even more important than the core – periphery exploitation (for example, without taking this mechanism into consideration it appears impossible to account for such things as the demographic explosion in the 20th century, whose proximate cause was the dramatic decline of mortality, but whose main ultimate cause was the diffusion of innovations produced almost exclusively within the World System core). This also suggests a redefinition of the World System (WS) core. The core is not the WS zone, which exploits other zones, but rather the WS core is the zone with the highest innovation donor/recipient (D/R) ratio, the principal innovation donor.[12]

[12] Earlier we regarded an "information network" as a sufficient condition to consider the entity covered by it as a "world-system". However, some examples seem to be rather telling in this respect. *E.g.*, Gudmund Hatt (1949: 104) found evidence on not fewer than 60 Japanese ships accidentally brought by the Kurosio and North Pacific currents to the New World coast between 1617 and 1876. Against this background it appears remarkable that the "Japanese [mythology] hardly contains any motifs that are not found in America (which was noticed by Levi-Strauss long ago)" (Berezkin 2002: 290–1). Already this fact does not make it possible to exclude entirely the possibility of some information finding its way to the New World from the Old World in the pre-Columbian era, information that could even influence the evolution of some Amerindian mythologies. However, we do not think this is sufficient to consider the New World as a part of the pre-Columbian World System. The Japanese might have even told Amerindians about such wonderful animals as horses, or cows (and some scholars even claim that a few pre-Columbian Amerindian images depict Old World animals [von Heine-Geldern 1964; Kazankov 2006]); the Japanese fishermen might even have had some idea of say, horse breeding. But all such information would have been entirely useless without some specific matter – actual horses or cows. Hence, now we would denote respective "system-creating" networks as "innovation diffusion networks" rather than just "information networks".