Mapping patterns of long-term settlement in Northern Mesopotamia at a large scale

Bjoern H. Menze[a,b,1] and Jason A. Ur[a]

[a]Department of Anthropology, Harvard University, Cambridge, MA 02138; and [b]Computer Science and Artificial Intelligence Laboratory, Massachusetts Institute of Technology, Cambridge, MA 02139

Edited by Frank Hole, Yale University, New Haven, CT, and approved December 23, 2011 (received for review September 21, 2011)

The landscapes of the Near East show both the first settlements and the longest trajectories of settlement systems. Mounding is a characteristic property of these settlement sites, resulting from millennia of continuing settlement activity at distinguished places. So far, however, this defining feature of ancient settlements has not received much attention, or even been the subject of systematic evaluation. We propose a remote sensing approach for comprehensively mapping the pattern of human settlement at large scale and establish the largest archaeological record for a landscape in Mesopotamia, mapping about 14,000 settlement sites—spanning eight millennia—at 15-m resolution in a 23,000-km^2 area in northeastern Syria. To map both low- and high-mounded places—the latter of which are often referred to as "tells"—we develop a strategy for detecting anthrosols in time series of multispectral satellite images and measure the volume of settlement sites in a digital elevation model. Using this volume as a proxy to continued occupation, we find a dependency of the long-term attractiveness of a site on local water availability, but also a strong relation to the relevance within a basin-wide exchange network that we can infer from our record and third millennium B.C. intersite routes visible on the ground until recent times. We believe it is possible to establish a nearly comprehensive map of human settlements in the fluvial plains of northern Mesopotamia and beyond, and site volume may be a key quantity to uncover long-term trends in human settlement activity from such a record.

Assessment of the scale and spatial distribution of human communities in past societies has been a key objective for archaeological research over the last 50 y (1), especially in the Near East, where questions of the origins of urbanism, the state, and empires are almost always approached via a regional perspective (2), and most commonly using the methods of archaeological survey. The pioneering surveys of the 20th century focused on the top of the settlement hierarchy, in the form of the largest mounds (e.g., refs. 3 and 4), but it is now appreciated that by overlooking smaller sites such an approach can produce misleading portraits of settlement systems, particularly those of nonurbanized phases (5). The challenge for archaeologists is to maintain coverage extensive enough to discern significant spatial patterning while increasing survey intensity in order to locate these smaller sites. The first generation of survey archaeologists opted for the former half of this classic tradeoff; more recent generations have increasingly adopted the latter intensive approach.

Many surveys have met this challenge by employing various remote sensing datasets. Of the three primary physical properties characterizing sedentary Near Eastern sites since the Neolithic—dense surface artifact assemblage, moundedness, and anthropogenic sediments—the latter two can be detected in aerial or satellite imagery and can potentially be recorded at large scale. Indeed, nearly all Near Eastern surveys have employed some form of satellite remote sensing, although usually at coarse resolution (6, 7) or for very limited study areas (8, 9) only, and nearly always in a qualitative fashion (6, 8, 9). Here, we propose a remote sensing approach that is simultaneously extensive and intensive, and which exploits two of the defining features of ancient settlement sites in the Near East: The proposed approach recognizes anthropogenic sediments via scenes from multispectral satellite images in a comprehensive fashion and measures moundedness of archaeological sites via a digital elevation model.

By relying on site volume, we also systematically survey an underappreciated feature of Near Eastern settlements: the mounding of the site. Surprisingly, the volume of settlement mounds has attracted little interest in the analysis of settlement systems, up to the present. Volume, however, is a function of the population present at a site through times and may serve as a quantitative measure for identifying the environmental or economic conditions under which settlements of agrarian societies in this part of Mesopotamia may have flourished. We propose to use this record of the debris present at a site as a proxy measure for quantifying the long-term attractiveness of settlement locations, and for ranking them in accordance to their ability to attract and sustain human communities over long periods of time.

Mapping Anthrosols at a Large Scale

In the alluvial and largely treeless plains of the northern arc of the Fertile Crescent (10)—as in many other landscapes of the Near East—the primary building material was mud brick. As dwellings became dilapidated, their walls were partially dismantled and flattened, and new structures were built atop their remains. In this manner, settlements grew vertically in the centuries or millennia that they were occupied (11), and soils at the site transformed to characteristic anthrosols. The largest mounds—known as tells in Arabic—still rise to heights of dozens of meters and represent well-known landmarks in many regions of the Near East.

The Upper Khabur Basin of Hassake Province, northeastern Syria (Fig. 1), is known as a "landscape of tells" (12, p. 100), extending about 200 km in east–west direction and 100 km from north to south. This basin in northern Mesopotamia is a critical locus for the study of the origins and development of urbanism in the Near East, and for the study of the organization of settlement under early states and empires (see reviews in refs. 13–15). Therefore, we choose this region to develop and test strategies for mapping anthrosols at large scale and for quantifying the volume of all sites visible from this map.

Localizing Anthrosols in Multispectral Satellite Imagery. The presence of anthropogenic soils is a defining feature of former settlement sites in large parts of the Near East. During millennia of human occupation, these anthrosols formed from organic waste and the decay of mud-brick architecture. Their texture, hydrological, and reflective properties often differ significantly from the surround-

Author contributions: B.H.M. and J.A.U. designed research; B.H.M. and J.A.U. performed research; B.H.M. contributed analytic tools; B.H.M. and J.A.U. analyzed data; and B.H.M. and J.A.U. wrote the paper.

The authors declare no conflict of interest.

This article is a PNAS Direct Submission.

Data deposition: Anthrosol map and derived data products can be downloaded from the Harvard Institute for Quantitative Social Sciences data repository via http://hdl.handle.net/1902.1/17731.

[1]To whom correspondence should be addressed. E-mail: menze@csail.mit.edu.

See Author Summary on page 5146 (volume 109, number 14).

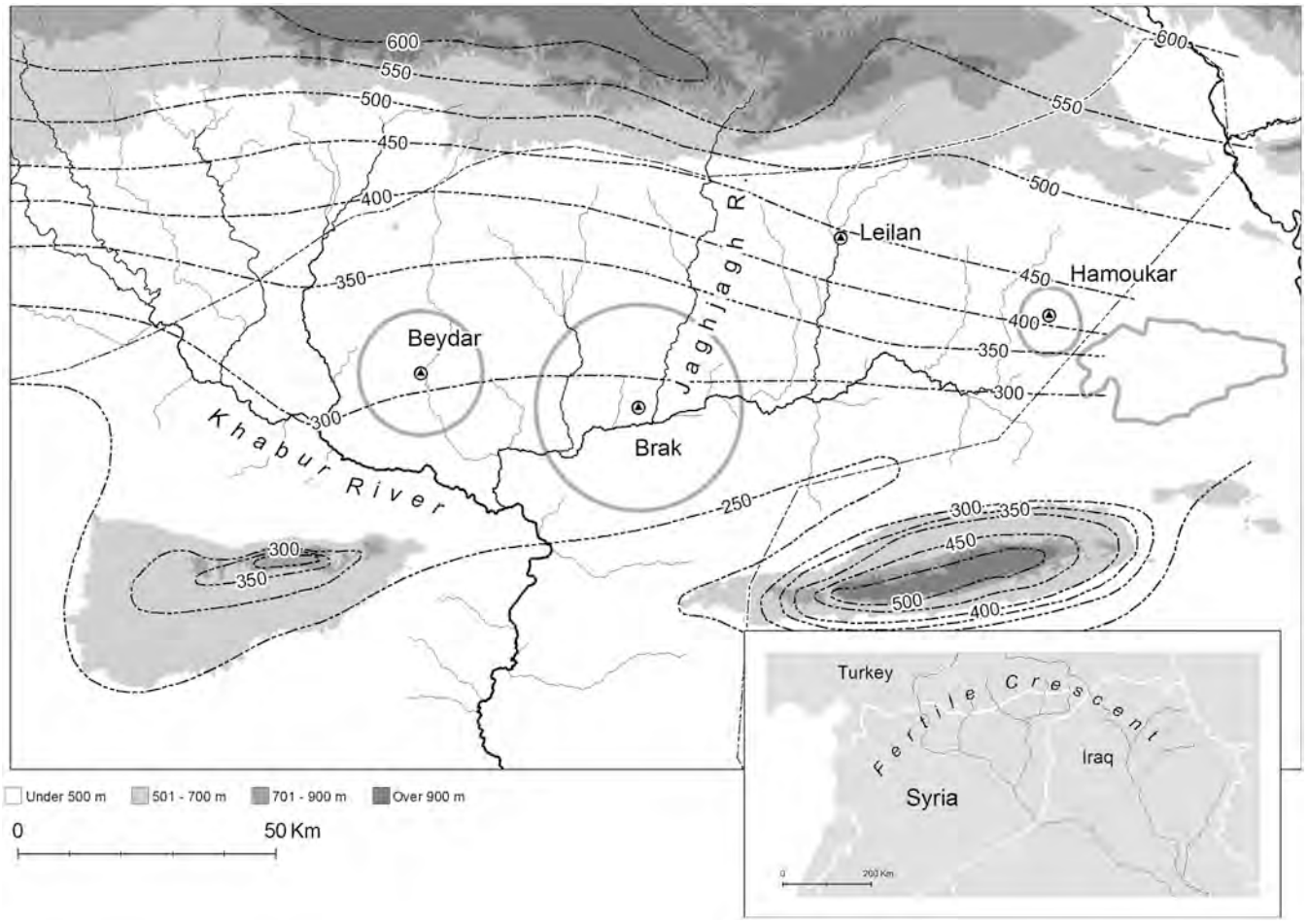

Fig. 1. The Upper Khabur Basin with rainfall isohyets and major archaeological surveys indicated; from west to east: Tell Beydar Survey, Tell Brak Survey, Tell Hamoukar Survey, North Jazira Project. (*Inset*) Localization in Hassake Province, Northeastern Syria.

ing soils, because they have a higher amount of organic components, a finer texture, and a lighter appearance (16). It makes them visible not only on the ground in regional surveys, but also in satellite imagery. As a consequence, the mapping of early settlements using satellite imagery is a central element in the analysis of Near Eastern landscapes. In particular, imagery from the declassified CORONA reconnaissance missions has proven useful for detecting sites and landscape features (17) (Fig. 2). These grayscale panchromatic photographs were acquired in the visible range and provide spectacular views of northern Mesopotamian landscapes of the 1960s and 1970s prior to disturbance by modern development and agricultural expansion (18, 19). Imagery from the CORONA program is extensive but still somewhat limited in spatial and temporal extent; for many areas of the Near East, scenes from only one or two missions are available, and many of them may have been acquired under conditions that are suboptimal for archaeological site visibility.

This limitation in the availability of appropriate satellite imagery can be overcome in part by newer multispectral sensors that revisit areas frequently throughout the year, because the archaeological matrix of anthrosols is also visible in the infrared spectrum recorded in multispectral satellite images (8). Unfortunately, due to the high specificity of the multichannel signal, the signatures of the structures of interest and their surroundings change to a much larger extent during the year than in standard monochrome imagery. At the same time, characteristic patterns in the multispectral signatures are difficult to infer for human observers due to the high dimensionality of the data and a relatively low spatial resolution of the images. As a consequence, the interpretation of imagery from sensors such as Landsat or Advanced Spaceborne Thermal Emission and Reflection Radiometer (ASTER) has only complemented the visual inspection of high-resolution aerial imagery and the analysis of multispectral images has not had a significant impact on regional surveys conducted in the landscapes of the Near East, so far, although their widespread use has been anticipated since data from the first multispectral sensors were available several decades ago (20). We propose to overcome this limitation by relying on a multitemporal approach capable of using the information not only from one, but all multispectral images covering a region of interest.

A multitemporal classification procedure. Because site visibility is closely connected to ground conditions, it may be insufficient to rely on a single satellite image for a given area. In ASTER images, for example, the spectral signature of the anthrosols changes interannually and throughout the year, related to rainfall fluctuations and the agricultural cycle (Fig. 2, *Left*). This variability is a significant problem when training a machine learning algorithm that is using the spectral signal of a few known archaeological sites as input to predict the location of other settlements in the same region; areas with similar hydrological conditions or those that are partially covered by the same vegetation as the archaeological sites may show up as false positives, and the resulting probability maps will always have a large amount of random "noise" (Fig. 2, *Right*). For this reason, we propose a multitemporal classification strategy that uses multiple ASTER scenes for any given position. It highlights those sites that show a spectral

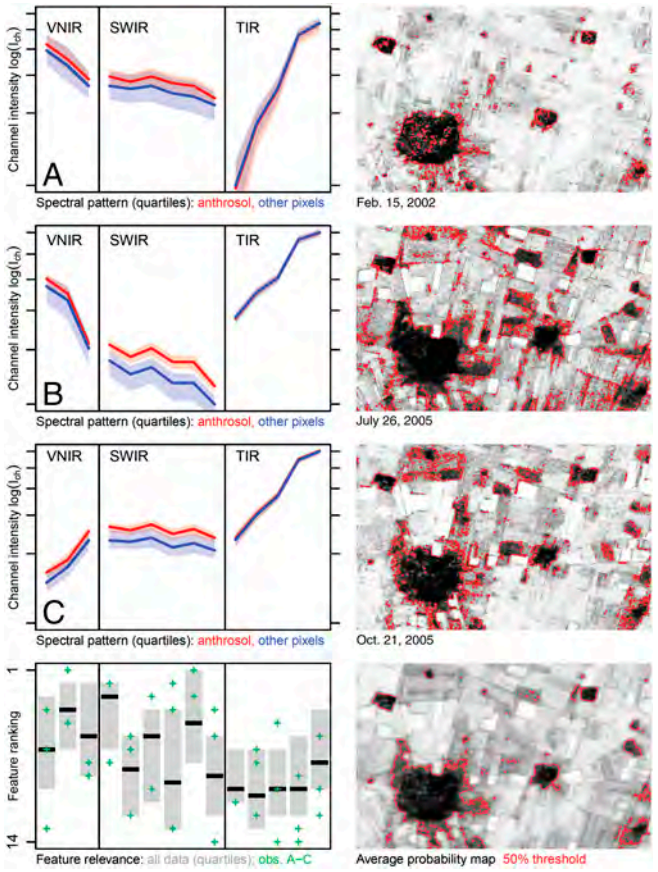

Fig. 2. Seasonal variation in spectral patterns and spatial probability maps for the Hamoukar region (Fig. 3*E*). Spectra for anthrosols (red) and other pixels (blue) are shown in the left column; lines represent medians, boxes quartiles (TIR has a different offset than the other channels; VNIR combines channels V and NIR). The anthrosol probability maps of observations *A*–*C* are shown in the right column (black 100%, white 0%; 50% threshold outlined red). The bottom row shows the feature ranking for all ASTER images of the study using the Gini importance (22, 24) as measure (boxes represent quartiles; black lines median; green crosses observations *A*–*C*). The final probability map—also shown in Fig. 3*E*—is obtained by averaging *A*–*C* and two more observations (which are similar to *B* and *C*).

pattern similar to that of nearby archaeological sites not just in a single image, but consistently over a long observational period.

To develop the classifier, we begin with 1,680 archaeological sites identified on CORONA images all throughout our search area; about 10% of them had previously been verified by ground survey (19). The image dataset comprises 160 multispectral ASTER images, acquired between 2003 and 2007 and spread evenly across all seasons; some parts of the survey area are covered by more than 40 scenes, whereas others have as few as four. (Because images are classified individually and only the resulting probability maps are fused, as described below, we could extend our dataset with scenes from other sensors, such as Landsat.) ASTER images have 14 spectral channels with reflectances from visible red (V) via near- and short-wave infrared (NIR, SWIR) to emissions in long-wave or thermal infrared (TIR); the spatial resolution differs from 15 (V) to 90 m (TIR). All observations were registered to a panchromatic Système Pour l'Observation de la Terre (SPOT) scene covering the whole basin at 10-m resolution using manually selected control points and then resampled to SPOT resolution using nearest-neighbor interpolation. The detection problem is then transformed in a binary classification task by amending the known archaeological sites with a second group of randomly chosen locations that are assumed to be non-archaeological sites. For every pixel at the locations of both the "settlement" and "nonsettlement" class, a number of descriptors of the spectral information are recorded: the original reflectances from the 14 ASTER channels, three standard vegetation indices (NIR/V; NIR – V; (NIR – V)/(NIR + V)) representing differently normalized reflectances, and the correlation of the full spectrum with different template spectra from the "ASTER SpecLib" representing soil types, water, minerals, and vegetation (21).

Using these descriptors for settlement and nonsettlement locations within the field of view of an ASTER image, we train a "random forest" classification algorithm to discriminate between the two classes (22). This nonlinear ensemble classifier performs an internal feature selection, choosing features with a high relevance to the high-dimensional classification task while safely ignoring irrelevant predictors. Other classifiers with similar properties, such as "oblique random forests" representing ensembles of multivariate decision trees and outperforming standard random forests on correlated features (23), could be used here as well. For any random forest classifier, the feature selection can be visualized. Using the "Gini importance" (24) to rank features over all ASTER images of this study (Fig. 2, *Bottom Left*), we find the information of ASTER channels 3, 5, and 6 to be the most important during training. After training, when presented with the spectral descriptors of other pixels from the same image, the classifier returns a probability indicating how likely anthrosols of the settlement class are present at those locations—i.e., continuous values which are close to zero for nonsettlement pixels and close to one for anthrosols and settlement sites. The random forest classifier requires little parameter optimization during training, and we adapt the algorithm individually to every satellite image available for our study region. We repeat this procedure in a spatially blocked cross-validation to guarantee a fair test for sites that are part of the training dataset (25). To this end, every available ASTER image is subdivided into patches ("blocks") of 6 × 6 km, which have a side length well above the average correlation length of archaeological feature on the ground. The random forest classifier is then trained on data from all patches but one and is used to predict the location of archaeological sites in the one hold-out patch. This procedure is repeated over all 36 patches within each scene. As a positive side effect, this cross-validation also reduces the influence of false positive or false negative training labels on the outcome of the classification: Training sites of the settlement class that do not share the spectral characteristics of other settlements in their surroundings do not show up in the final anthrosol probability map.

After all ASTER images are classified, the resulting probability maps are fused (Fig. 2, *Right*), producing an average signal that is largely free of noise from short-term variations of the spectral signal, and highlighting only those locations that *consistently* show the same spectral signature as nearby archaeological sites in the settlement class. As the average probability still fluctuates somewhat based on soil, soil cover, and how those changes are represented in the local training data, the probability map is represented as a grayscale map and inspected visually in a manner similar to, for example, a visual inspection of a CORONA image (Fig. 2*A*). The most obvious sites—i.e., those with characteristic shapes in the probability map or with further signs of archaeological sites in high-resolution satellite images (SPOT, IKONOS)—are identified, and the training dataset is amended to include these sites for a better spatial coverage, before the whole classification process is repeated once again. After two iterations of this procedure, the final evaluation of the map returns a total of 14,312 sites. These sites stand out from the spectral signature of their surroundings and typically show the round morphology of early settlement sites. An automated routine determines the extensions of every recorded site and measures its area in the anthrosol map. In this analysis, about 12,000 sites are larger than 1 ha, and a total area of 856 km^2 displays signs of ancient or present settlement, which amounts to about 4% of the total search area. Out of these sites, about 5,350 (or 461 km^2) show further

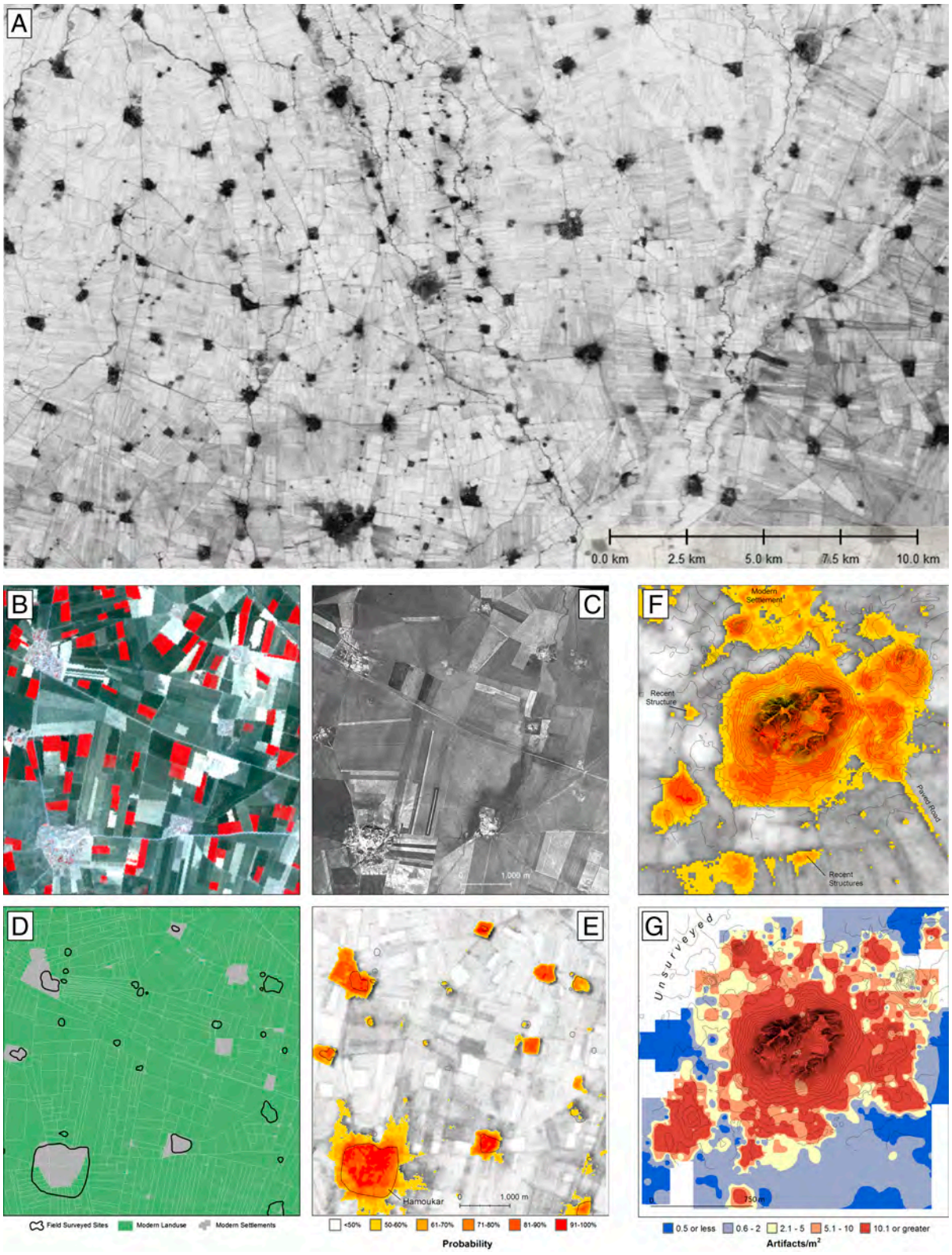

Fig. 3. Anthrosol probability map (*A*) and validation for the Hamoukar (*B–E*) and Brak region (*F* and *G*). (*A*) Anthrosol map for the east-central part of the basin; dark pixels indicate high probabilities. The large site in the northeastern corner is Tell Leilan (Fig. 1). (*B*) ASTER image for the Hamoukar region from October 18, 2004 (bands 3N, 2, 1). (*C*) CORONA image 1108, December 6, 1969. (*D*) Sites, modern villages, and contemporary land use from field observations, 1999–2001. (*E*) Anthrosols probability derived from multitemporal ASTER analysis (see Fig. 2). (*F*) Anthrosols at Tell Brak, the largest mound of the basin (same scale legend as in *E*). (*G*) Sherd density recorded in the Tell Brak Suburban Survey.

signs of anthrosols or early settlements when examined in high-resolution SPOT or CORONA images. To validate results, we compared them against field data of several archaeological surveys in the region (Fig. 1) and against other proxies of long-term settlement activity with ground truth available at the pixel level.

Validation against ground survey. We have detailed knowledge about sites in the hinterland of Tell Hamoukar (Fig. 3 *B–E*), a 100-ha Early Bronze Age (*ca.* 2600–2000 B.C.) urban site in the eastern part of the basin (Fig. 1). A survey of 125 km^2 surrounding Hamoukar recorded 60 sites, most of them being first identified in CORONA and then validated on the ground (19). The classification algorithm identifies 132 sites, including 52 of the sites recognized in the archaeological survey (87%). The eight field sites missed by the analysis are all flat or nearly flat, and most had been subjected to aggressive agricultural modification

in the last decades (i.e., in between the acquisition times of the CORONA images in the 1960s and the ASTER images 35 y later). Many had been plowed deeply or bulldozed or had been heavily irrigated; none are higher than 1 m or larger than 2 ha. At the same time, other sites which had been subjected to similar transformation were detected successfully. Of the 82 additional sites identified by the anthrosols classifier, 37 are modern villages or other areas of recent human disturbance (Fig. 3 *D* and *E*). The remaining 43 sites comprise 28 places that might be sites missed by the field survey (6.35 ha in total) and 15 places that appear to be likely false positives (3.97 ha in total).

Similar validation is found in other surveyed regions (Fig. 1*A*). In the western part of the basin, the area around Tell Beydar (452 km^2) contains 82 settlements (26, 27); 60 of them are clearly visible by their anthrosols, and 11 further sites show a weak signal (73–87% detection rate). Most of the sites that are not visible in the anthrosol map are smaller than 1 ha. About 50 other locations show a high probability in the anthrosol map, with many (but not all of them) representing modern settlements. The regional survey around Tell Brak, the largest mound in the basin, reports a total of 268 sites (as of 2003) (28). Of these, we find 249 sites with a strong anthrosol signal, and a weak signal for 10 of the remaining sites, all of which are rather small (93–97% detection rate). Around Tell Leilan, in the central part of the basin, an area of approximately 1,600 km^2 had been surveyed (29, 30). Out of 300 recorded and published sites, 286 show a high probability, and another three show a weak one (95–97% detection rate). Of the 11 sites missed, seven are small sites close to Tell Leilan, presumably the most frequently visited area of this survey. In the southeastern end of our study region, the probability map partially overlaps with the North Jazira Project survey in northern Iraq (517 km^2) (31). Here, 77 of 100 former settlement sites are clearly visible, and another 10 showed an elevated anthrosol probability (77–87% detection rate). These results are remarkable because only four ASTER images are available for this area, and—as a rule—better results are obtained with more images.

Although overall 90–94% of the sites identified on the ground are also visible from the anthrosol map (672–706 out of 750 sites in total)—and many additional likely sites can be identified—the number of false positives is more difficult to quantify. First, ground truth might be somewhat unreliable, depending on the intensity of archaeological survey and how each survey defined their sites. Second, small modern settlements appear in the anthrosol map with high probability. Sometimes they sit atop ancient mounds, but often they just show a spectral signature similar to the one of early settlements in their surroundings. In most cases, however, these "false positive" modern villages can be recognized through visual inspection of standard high-resolution imagery and the measurement of site volume (described below).

In the special case of Tell Brak (Fig. 3 *F* and *G*), one of the largest Bronze Age sites in northern Mesopotamia, it is possible to evaluate the results of the anthropogenic soil classification against the one archaeological indicator of human occupation that cannot directly be detected from satellite images: surface artifact density. An intensive field survey of the surface of Tell Brak systematically recorded sherd densities at 50- to 100-m intervals (Fig. 3*G*), in an area about 5 km^2 in size (32, 33), and one can evaluate true and false positives even at pixel level (Fig. 3*F*). When comparing the interpolated sherd density and high probability of anthropogenic soils, the site extends correlate almost precisely. These results confirm that the anthrosol map not only identifies archaeological sites, but also that it provides a reliable estimate of sites' shapes and extents, revealing—for example—the border of lower towns around high-mounded sites, which otherwise require intensive field surveys.

Measuring Site Volume in a Digital Elevation Model. In addition to the presence of anthropogenic soils and sherds, mounding is the third defining feature of former settlements. With the release of the global Shuttle Radar Topography Mission (SRTM) digital elevation model, archaeologists gained a valuable tool for recognizing mounded sites (6), and the authors have developed an automated method for mapping them at a broad regional scale (7), although only the largest mounds can be identified from the elevation model alone. Using the anthrosol map as guidance for small and medium-sized mounds, however, it may be used to estimate volume of these sites, too.

We use the 90-m resolution SRTM digital elevation model (DEM) dataset and interpolate it to the 10-m resolution of the anthrosol map (nearest-neighbor interpolation). We then estimate the plain level upon which the settlement was initially constructed, relying on a bivariate spline model adapted to the terrain of the surrounding of the site using randomly chosen support points. This basal level is subtracted from the observed elevation, and the volume is calculated from those pixels in the center of the site that stand out significantly from the residual DEM (by more than two standard deviations of the residual in the surrounding topography, usually about 1 m). This calculation leads to an estimate of the volume of the central mound. Sites that have little or no volume—typically representing briefly occupied modern settlements, the most frequent false positives from the ASTER analysis—can be excluded from further analysis. Performing the volume estimation twice for randomly sampled sites throughout the basin, we find that volume estimates vary less than 10% between repeats (with the variation mainly resulting from differently initiated estimations of the basal level in areas with rugged terrain).

The validation of the estimated values remains challenging because the volume of settlement sites has never been recorded systematically in our region of interest (nor, to the best of our knowledge, in any other regional survey in the Near East), and the volume of a mound is reported only for select sites in an anecdotal fashion. Approximate estimates of volume V, however, can be determined from accurate measurement of area A and maximal height h of a mound, quantities recorded in some surveys (11, 34). The shape of a mound is in between the shapes of a short cylinder with volume $V_{cyl} = A \cdot h$ and an irregular shaped cone, or pyramid, with volume $V_{cone} = \frac{1}{3} \cdot A \cdot h$. In our study region, measurement of area and height are available from the field data of the Tell Hamoukar survey (19). When compared, approximated ground truth and volume from the DEM are in reasonable agreement over several orders of magnitude (Fig. 4). Observed differences can be explained by shape variations of the mound: Smaller sites are flat, cylinder-shaped mounds and the true volume will be at the upper end of the expected range, at around V_{cyl}, indicated by the upper hinges in Fig. 4, which is in good agreement with our DEM measurements for small sites. Large tells show pyramidal or even concave structures with a very sharp peak. For these mounds, the true volume would even be below the estimate for V_{cone} (i.e., below the lower hinges in Fig. 4). This is the position in the scatterplot where we find most of the DEM measurements for the largest sites.

Overall, 14,312 sites show the characteristic signature of anthrosols in the infrared part of the electromagnetic spectrum. A total of 9,529 sites elevate significantly from the ground, representing an area of 157 km^2. Whereas the average mounded area is 1.7 ha, there are 2.5% of the mounds larger than 10 ha, with an integral volume of 266 million m^3. This volume is a significant part of the total volume of all sites, which is 700 million m^3 when integrating over the central mound of the site ("high town," as described above) and 940 million m^3 when integrating over adjacent areas with anthropogenic soils as well ("lower town"). The largest individual mound in the basin is Tell Brak, which is estimated at 7,960,000 m^3; this number agrees well with the estimates of the excavators, who put it between 7 million and 8,890,000 m^3 (35). Among other sites, for example, the major third millennium B.C. urban center Tell Hamoukar has a volume

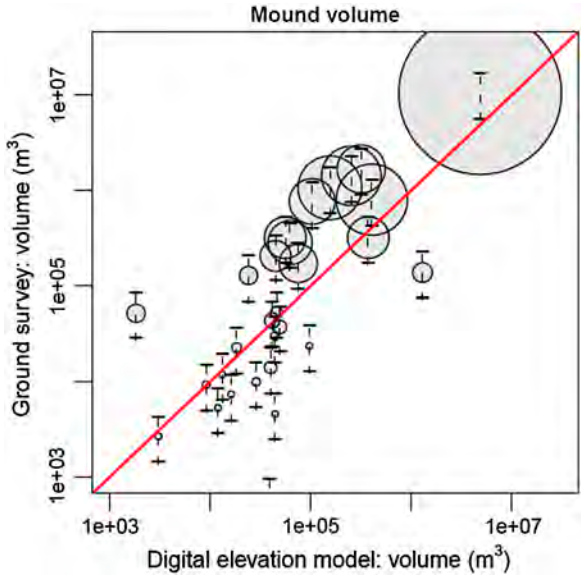

Fig. 4. Site volume. Estimates from the SRTM digital elevation model (*x* axis) and from field measurements for mounds in the Tell Hamoukar Survey region (*y* axis); vertical bars represent variation between conical and cylindrical volume models (also see text); circle size is a qualitative indicator of the area of the mound as measured on the ground.

of 4,660,000 m^3, and the third millennium town at Tell Beydar in the western part of the basin has a volume of 1,150,000 m^3. As a comparison, the debris of a single-household mud-brick building occupied for about one generation is estimated to be in the range of 20–100 m^3 for architecture similar to the one in the Khabur Basin (34). These numbers may provide a coarse, indirect measure of the total premodern sedentary population in the basin which—under the assumptions from ref. 34—would amount to 10–40 million households.

Patterns of Long-Term Settlement

This accumulated anthropogenic debris represents the aggregate of approximately 8,000 y of sedentary agricultural settlement. Nucleated and mounded settlements evolved along this northern fringe of the Fertile Crescent initially during the Neolithic period, *ca.* 7,000 B.C., and developed into the landscape of tells we now find in the Upper Khabur Basin. Unfortunately, analysis of recorded volume cannot suggest *when* a site was occupied, and chronological information is only available for those few sites that have been excavated or surveyed on the ground. As a whole, however, the spatial distribution of thousands of sites provides a quantitative basis for testing relations to different environmental factors that may have shaped the observed long-term settlement pattern. With volume as a quantitative indicator of long-term attractiveness of a settlement location at hand, we want to identify factors that may have been critical in shaping the observed spatial patterns of past populations.

In our study region, the environment displays few major topographic constraints except the two perennial rivers—the Khabur River and its tributary, the Jaghjagh (Fig. 1)—and a north–south gradient in precipitation and some variation in natural soil characteristics (36). For agricultural communities, availability of water and the local soil types are among the most critical variables for assessing the attractiveness of a settlement site (37). We correlate the integral settlement volume per area—as a proxy for the total population over time—with soil types, current precipitation (assuming its north–south gradient has been stable over time, though the absolute amount of rainfall may have fluctuated), and the catchment size of the largest runoff next to a settlement (as a measure for long-term availability of surface water; Fig. 5 *A* and *B*).

Correlation with Local Resources. We first test the relationship between settlement volume and soils. Although available soil maps differ slightly, only few soil types dominate the basin (38, 39). We aggregate the total site volume over local 10 × 10 km patches and associate volumes of all 232 such patches with the major soil type present in that area. For every soil type, the distribution of the observed volumes is then compared against the distribution of the most common soil (Fig. 5*C*). In this test, no significant differences in the observed volumes can be found for any of the major soil types ($p > 0.1$, Cox–Wilcoxon test). Only soils in the periphery of the basin, and close to the surrounding mountain ranges, show a significantly smaller integral settlement volume ($p < 0.01$), suggesting that soil differences had—at large scale—rather little effect on the overall distribution of site volumes in the basin.

So, we focus on relations between settlement volumes and availability of water, using present rainfall and surface drainage as proxies to earlier patterns. Under present climatic conditions, rainfall patterns in the Upper Khabur Basin fluctuate from north to south (Fig. 1). The northern fringes of the basin receive sufficient rainfall for successful dry-farming practices in most years;

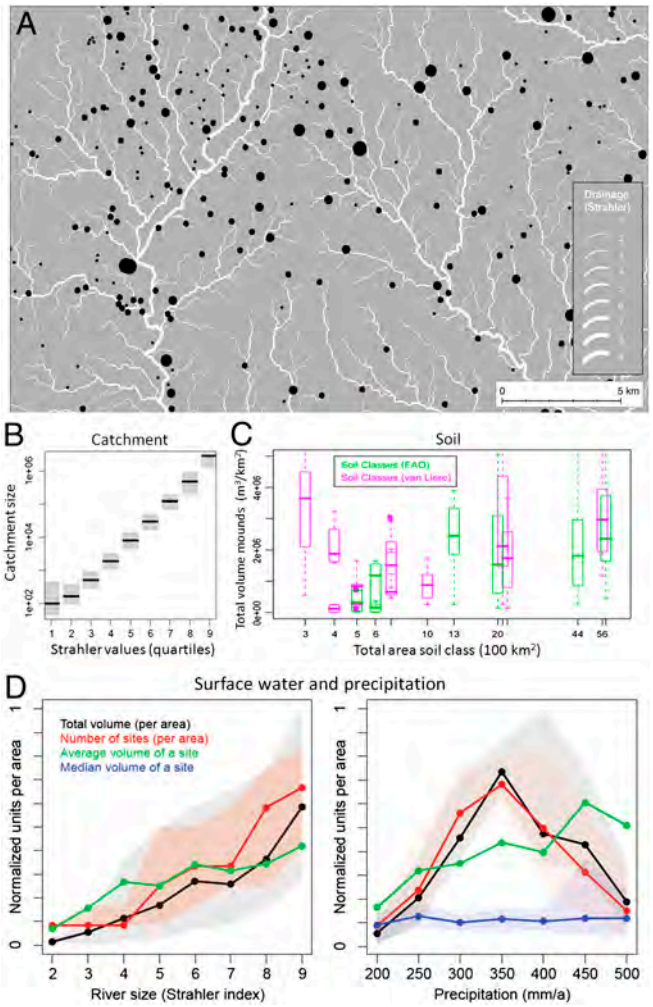

Fig. 5. Correlation of site volume and local resources. (*A*) Modeled surface drainage and sites in the hinterland of Tell Beydar (compare with streams shown in Fig. 1). (*B*) Log-linear relation between Strahler value and catchments area (boxes indicate quartiles, black line the median). (*C*) Relation between settlement density and soil type; box plots show the distribution of observed volumes for different soils (*y* axis), ranked with respect to the total area of the specific soil region (*x* axis). (*D*) Relation of settlement density with water availability; shown are median (lines) and quartiles (shaded areas). Units are normalized for visualization purposes.

the central plain is more risky, with a failed crop every 3 to 4 y, with even greater risk along its southern fringes (40). One may assume that the north-to-south precipitation gradient has been stable over long times, even when the absolute amount of rainfall may have varied. In that case, every 10 × 10 km area can be associated with the current precipitation rate—between 200 and 500 mm/y—and relative differences in modern rainfall may serve as an indicator to earlier precipitation differences and shifts of the dry-farming limit along the gradient. The availability of surface water, however, depends on the absolute amount of precipitation. At present, the Khabur and Jaghjagh rivers flow perennially and all other drainages (wadis) flow seasonally, but river flow may have varied in the past (40), so we use drainage size as a more reliable indicator for the water collection potential of a stream and surface water availability for nearby settlements instead (Fig. 5 *A* and *B*). This descriptor, the Strahler value, is generated from the SRTM-derived network of surface drainages. It assigns a stream order of one to the smallest drainages; at the confluence of two first-order drainages, a second-order drainage originates, and so forth. For the Khabur Basin, this classification results in stream order values between one and nine, in excellent correlation with catchment size (Fig. 5*B*). Every site is associated with the largest stream value within 1,000 m to make it robust against past local changes of river channels and every 10 × 10 km patch with the largest river indicator in its area. Both quantities, local precipitation differences and catchment size of the local river, should be seen as indicators of a temporally integrated binary statistic: Precipitation differences correlate with the number of seasons of sufficient precipitation for crop production, and size of the nearby stream correlates with the number of seasons this river had met the local demand of water for a settlement, or had not.

Using these two indicators of water availability, we seek for correlations with three different measures of volumes: the number of sites per area, the integrated volume of all sites per area, and the average volume of the sites per area (which is the ratio of the first two measures) (Fig. 5*D*). All three quantities increase dramatically with drainage classification (*Left*). The number and the total volume of sites peak—somewhat surprisingly—at an annual precipitation value that corresponds to the middle of the plain, whereas the volume of individual sites increases with precipitation (*Right*). This trend applies to the average volume per area but not to the median, suggesting that higher precipitation has allowed a few sites to grow large. We seek for more simple rules to explain these observations.

A Resource-Dependent Long-Term Settlement Potential. To establish an average trend between the recorded quantities, we group all 14,312 sites according to local precipitation and size of nearby drainage, and we calculate the average volume for every combination of those variables (Fig. 6*A*). The resulting bivariate histogram (*Left*) shows a very systematic pattern: Volume correlates with precipitation (*Center*) and local drainage size (*Right*). An additive model with nonparametric terms (Fig. 6*B*), close to a log-linear additive model

$$\log(\text{site volume}) = c_1 \cdot \text{precipitation} + c_2 \cdot \text{river size} \qquad [1]$$

with constants $c_1 = 0.0018/(\text{mL/y})$ (*Center*) and $c_2 = 0.0737$ (*Right*), explains the log-transformed volume of a site in an excellent approximation to the observations (*Left*).

This model represents a strict exponential relation between integrated site volume and local water availability, indicating that small differences in precipitation and surface water may have led —and still lead—to drastic differences in the long-term attractiveness of a site. Interestingly, the two water variables are additive, suggesting that a lack of one water source can be offset by an excess of the other. This exchangeability may also apply to irrigation—compensating for a lack of precipitation or river water in the same way—and emphasizes the potential relevance of irrigation for maintaining settled places over long times.

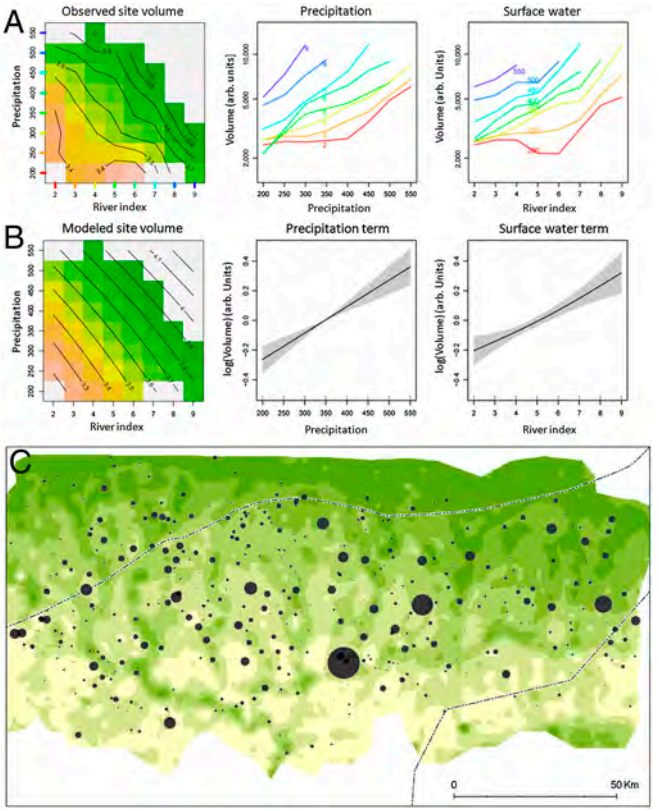

Fig. 6. Additive model defining settlement volume by precipitation and river catchment size. (*A*) Bivariate histogram (*Left*) of the average size of a site with the given precipitation and river value [yellow indicates low values ("dry"), green high ("wet"), gray are not observed], and volume as a function of precipitation (*Center*) and as a function of surface water (*Right*) with the same colors as indicated in the bivariate histogram. (*B*) Average size of a site as estimated in the additive model (*Left*), precipitation term of the additive model (*Center*), and river size term (*Right*), correlating well with observations from above; both additive terms are nearly linear, gray areas indicate 90% intervals from cross-validation; dotted lines in the background correspond to the linear model terms reported in the text. (*C*) Estimated settlement potential using the model in *B*, with settlements exceeding the potential indicated by dots.

We can use those functional relations to map the average long-term settlement "potential"—i.e., the expected volume of a site under the given resources—at any given location in the basin. To do so, the water index and precipitation is determined for 50,000 randomly chosen locations, and the expected size of a settlement at any of these locations is estimated using the additive model (Eq. **1**) from above. In the map obtained when spatially smoothing the predicted values (Fig. 6*C*), the floodplains of the Khabur River and its tributaries stand out as high potential zones running through otherwise low potential areas in the southern part of the plain. Most of the large sites are within or in direct vicinity of regions standing out as "islands" from their surroundings in the potential map. A different situation pertains in the northern half of the basin, where the long-term settlement potential is relatively homogeneous and appears to be determined largely by precipitation. The wider availability of "high potential" land in the northern and central areas and the necessity to build settlements at few attractive sites in the south may explain some of the features of Fig. 5*D*. In the southern half of the basin, steep gradients in this potential lead to a concentration on those few sites close to the most favorable river. This observation is in accordance with few dominating sites in the south, such as Tell Brak, and rather few but stable and large sites in the north, which could develop under

stable conditions over longer times. In the middle of these extremes, sites are repeatedly settled—abandoned over time with vanishing precipitation, then being resettled once again—leading to a large number of sites in the central areas of the plain.

Beyond large-scale average patterns, at an individual site level, the potential map may also be used to identify sites which surpass the expected limit and stand out from the expected average (Fig. 6*C*, black dots). We observe a dense net of large sites in all areas of the basin which evolved in size far beyond the expectation defined by the local resources. Most sites are in a 5- to 10-km distance with 12- to 30-km^2 area of local support. These systematic deviations in long-term settlement preferences suggest cultural factors behind settlement choices.

A Regional Long-Term Exchange Network. Sustaining settlements at these preferred locations over long time could only be accomplished with the efforts of a group of people that were significantly larger than the population of the settlement—i.e., in a settlement system. Fortunately, we have basic understanding about an extensive premodern intersite transportation network in the Khabur Basin (18). This "hollow way" network formed during 2600–2000 B.C., coinciding with a time of intensive settlement activity in the Khabur Basin, and remained visible until modern times. We had mapped more than 6,000 km of hollow ways in a previous study (18, 19) using CORONA images (Figs. 7 and 8). Although movement to fields and pastures may have had a significant impact on the formation of the hollow ways (18), a large fraction of these tracts connects settlements directly. We hypothesize that many of them may have evolved together with the settlements in the plain and show a similar continuity as the populated places they were connecting.

Together with our knowledge about site locations, we can use this hollow way record to infer the structure of the complete network of the basin. Human interaction is frequently modeled by exponentially truncated power functions (41). Under this exchange function—and using both site locations and recorded mounding—we can formalize the interaction E_{AB} between sites A and B, with distance d_{AB}, by

$$E_{AB}(\alpha, \beta, \gamma) = V_A^{\gamma} \cdot V_B^{\gamma} \cdot d_{AB}^{-\alpha} \cdot \exp(-d_{AB}/\beta). \qquad [2]$$

Here, we use site volume V as a proxy to site population in A and B, assuming that exchange is proportional to the population at either site. We introduce the exponent γ to vary the influence of V on the overall exchange E. Choosing different values for γ may account, for example, for a systematic mismatch in our assumption of contemporary settlement activity in A and B, or may transform our volume estimate in a number proportional to the average settlement area ($\gamma = \frac{2}{3}$), which is a common estimator of a settlement population.

We test E for different parameterizations within $\alpha = [0, \ldots, 7]$, $\beta = [0.25, \ldots, 16]$ km, and $\gamma = [0.01, \ldots, 2]$. We assume that E corresponds to traffic on the ground and map this directed local "traffic" to every pixel in between the connected sites. At every

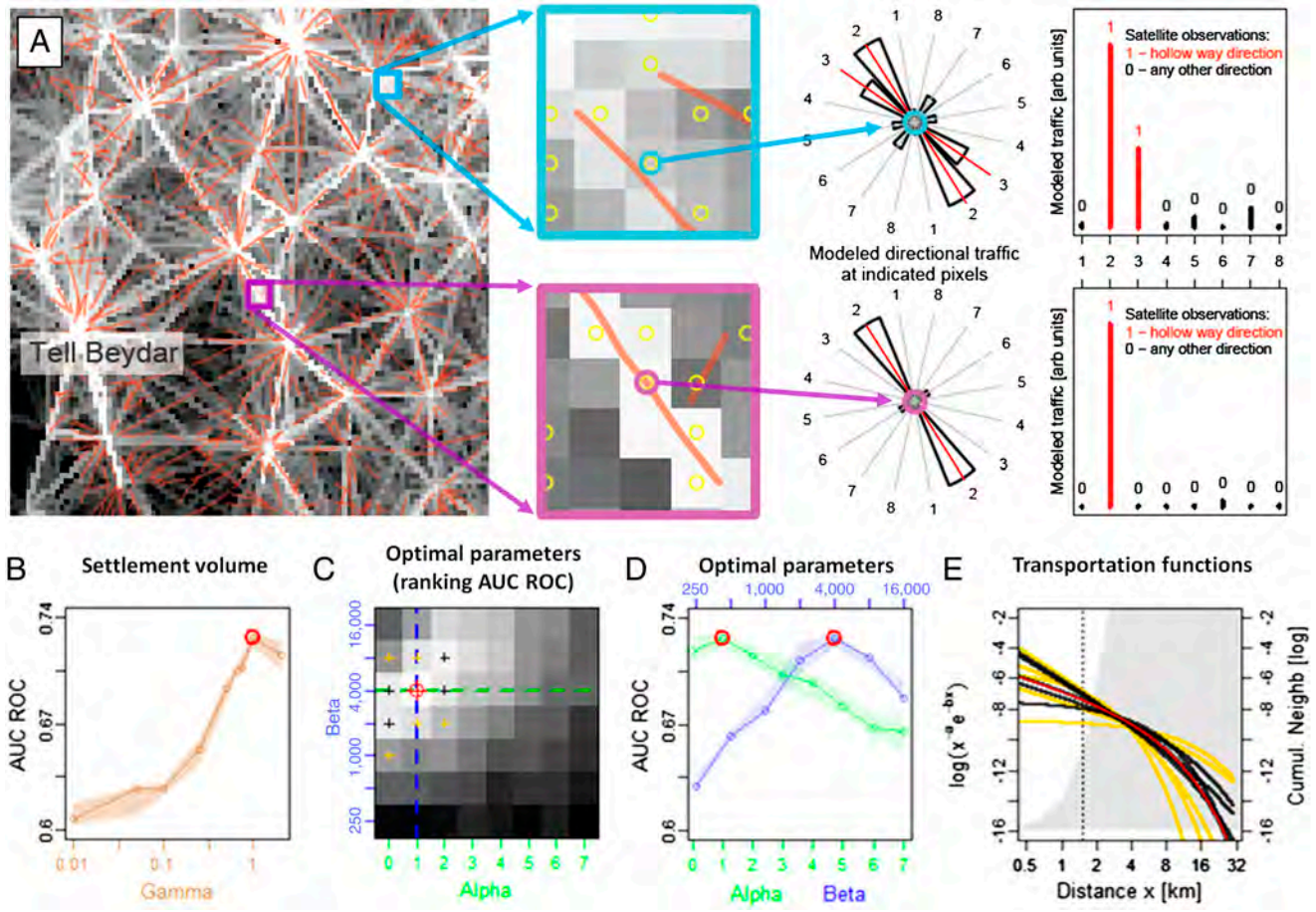

Fig. 7. Intersite connections. (*A*) Parameterization of the intersite exchange model with two examples from the Tell Beydar region (Fig. 8); goodness of fit is estimated at hollow way locations by the difference of traffic in hollow way direction (red, class "1") and any other direction (black, class "0"), measured by the area under the curve (AUC) of the receiver-operator-characteristic (ROC). (*B–E*) Results for different parameterizations of the exchange model; lines indicate median; shades indicate quartiles of 10 differently sampled datasets. Optimal values are indicated in red (*B* and *D*), although nearby parameters indicated black (crosses in *C*; functions in *E*) behave similarly well for distances beyond 1,500 m (dotted line, *E*), all differing visibly from functions and parameters indicated yellow. AUC ROC, area under the curve of the receiver operator characteristic.

hollow way location, we then measure the cumulative intersite traffic flowing in the direction of the hollow way, and the remaining traffic intersecting with the hollow way track at arbitrary orientations. Fig. 7*A* shows an example for a map with the absolute amount of traffic for the Tell Beydar area, together with the observed hollow ways (*Left*; at about 300-m resolution). Under an optimal parameterization of the transportation model, we expect the traffic that corresponds to the direction of the hollow way to be much larger than the remaining random undirected traffic (*Center*). We compare absolute values of both types of traffic (*Right*) at 2,000 randomly sampled hollow way locations using a statistical test measuring class separation (area under the curve of the receiver operator characteristic) and pool results over all test locations.

Results suggest (Fig. 7*B*) that volume parameter γ is the single most important parameter for explaining the observed patterns. A direct proportionality ($\gamma = 1$) of E and V performs best, whereas a transportation function relying on site locations but not site volume (with $\gamma \approx 0$) leads to nearly random results and cannot explain the observed hollow way patterns. The parameters of the distance function (α,β) are correlated (Fig. 7*C*), with optima around $\alpha = 1$ and $\beta = 4$ km (Fig. 7*D*). Different parameterizations perform similarly well (Fig. 7*C*, red and black crosses), although all top performing parameterizations—red and black functions in Fig. 7*E*, corresponding to crosses indicated in Fig. 7*C*—differ only beyond 1,500 m, a characteristic lower value for the distance of neighboring settlements. The observed maximum in the parameterization of the power function (with $1 < \alpha < 2$) is in good agreement with other patterns of human transportation (41, 42). The range of β—which has a characteristic length of 4–8 km—indicates that a significant amount of the traffic directly aimed at settlements beyond the immediate neighbors. Overall, the observed correlation of hollow ways structures with the reconstructed intersite exchange suggests a transportation network that is more complex than one solely arising from nearest-neighbor triangulations of settlement locations (43), and that considering site relevance—here through site volume V—is crucial for explaining the hollow way structures observed in between sites.

To understand the interaction between populated places and their long-term importance, we use the transportation function (Eq. **2**) with $\alpha = 2$, $\beta = 4$ km, $\gamma = 1$—in good agreement with the recorded hollow ways—to infer the complete network of the Khabur Basin (Fig. 8). Within this network, we measure site relevance as the number of visits a random traveler—or a certain good traded and passed on between sites—would make at a particular site of the network, a popular measure of network importance (44). Under this measure, sites with few connections to other infrequently visited places will show a low relevance, whereas sites with many connections to well connected hubs will receive a large number of visits. We find (Fig. 8, *Inset*) a strict relation between site volume and site importance within the overall network. As a consequence, in the hollow way network that emerged in the Upper Khabur Basin, it was not geographic centrality that determined the average number of visits a site received from intersite traffic in basin—as expected from a "location only" network (with $\gamma = 0$) as traditionally used in spatial analysis (43)—but rather it was the absolute size of the site.

Overall, site volume proves to be crucial for explaining the hollow ways in between sites and it is not site location, but absolute size that determines its relevance in the basin-wide network. As a consequence, large sites dominate both local and large-scale exchange in the regional network. This network property suggests

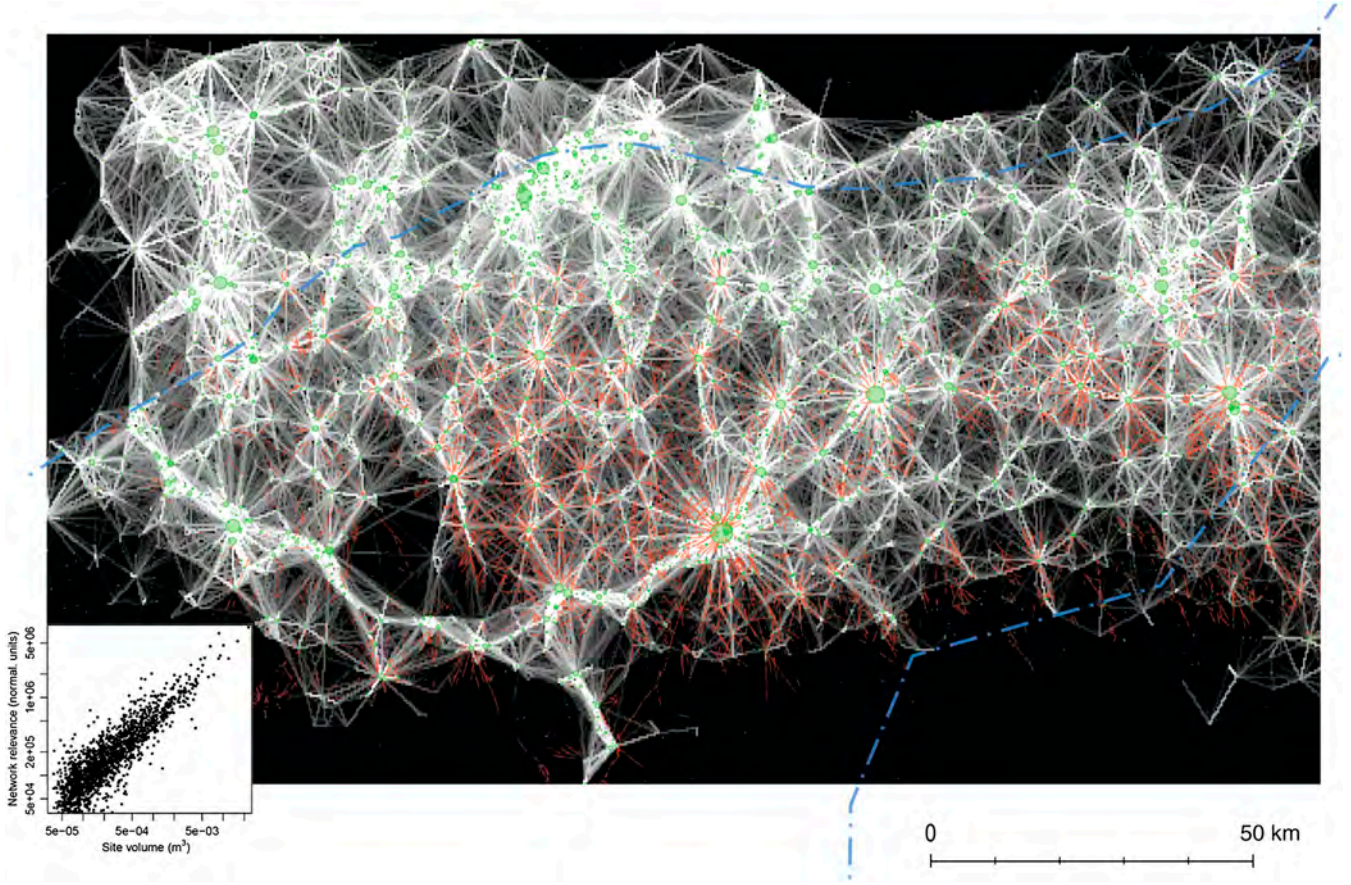

Fig. 8. Basin-wide network. Size of green dots indicates site volume and network importance; brightness of white lines indicates strength of modeled intersite connection; red lines represent hollow ways recorded from CORONA images (18, 19). The blue dashed lines represent modern political boundaries, also shown in Fig. 1. (*Inset*) Correlation of site volume with site relevance in the basin-wide network.

that local interaction in the settlement network is organized to have a stabilizing effect for those settlements whose growth exceeded the limit expected from local resources, helping to sustain such sites over centuries or millennia. From a wider perspective, this network property also suggests that patterns of movement would have been as stable as patterns of settlement.

It might be predicted that combined models of the local distribution of resources, mounding, and intersite connectivity would provide further insights into the organization of settlement hierarchies and regionalization within this basin-wide interaction network.

Conclusions and Prospects

We present a remote sensing approach that systematically maps anthrosols accumulated over eight millennia of settlement history, and establish the largest archaeological remote sensing record for ancient landscapes in Mesopotamia. Applied to the Upper Khabur Basin, a first test returned the locations of 14,312 sites over approximately 22,000 km^2, in total covering more than 3% of the search area. Of these, settlement mounds cumulate to an area of 157 km^2 and a volume of more than 700 million m^3 of collapsed architecture and other settlement debris, making it, to the best of our knowledge, the largest systematic satellite-imagery-based survey in archaeology.

Our multitemporal classification strategy can be easily generalized for other detection tasks in archaeological remote sensing, also integrating information from different multispectral sensors. The findings also stress an underappreciated feature of ancient Near Eastern settlement with great significance for long-term human occupation and the durability of a place's relevance: site volume. We demonstrate the significance of this proxy for describing generalized long-term patterns of human settlement and land use and illustrate its particular relevance as an indicator of past populations in spatial analysis.

We envision a nearly comprehensive map of sedentary human settlement for the fluvial plains of Northern Mesopotamia using the proposed approach, with results to be used in the integration of published findings, the planning of new surveys, and heritage management at international scale. Moreover, by measuring site volume, it may be possible to uncover long-term trends in human settlement activity from such a large-scale record, potentially revealing further environmental factors that govern long-term trends and sustainable settlement in the Near East.

ACKNOWLEDGMENTS. The authors acknowledge financial support from the Fritz-Thyssen-Stiftung and the German Academy of Sciences Leopoldina (through the Leopoldina Fellowship Programme, LPDS 2009-10; to B.H.M.), and further support by Simone Muehl and Peter Miglus, Heidelberg University. This work is dedicated to the late Andrew G. Sherratt whose contribution was crucial in initiating this study; read his account on first joint efforts toward the systematic "Tellspotting" presented here in the ArchAtlas case studies (45).

1. Kowalewski SA (2008) Regional settlement pattern studies. *J Archaeol Res* 16:225–285.
2. Wilkinson TJ (2000) Regional approaches to mesopotamian archaeology: The contribution of archaeological surveys. *J Archaeol Res* 8:219–267.
3. Adams RM (1981) *Heartland of Cities* (Univ of Chicago Press, Chicago).
4. Braidwood RJ (1937) *Mounds in the Plain of Antioch: An Archeological Survey* (Oriental Inst, Chicago).
5. Wilkinson TJ, Ur JA, Casana J (2004) From nucleation to dispersal: Trends in settlement pattern in the northern fertile crescent. *Side-by-Side Survey: Comparative Regional Studies in the Mediterranean World* (Oxbow Books, Oxford), pp 198–205.
6. Sherratt AG (2004) Spotting tells from space. *Antiquity* 78:301, http://antiquity.ac.uk/projgall/sherratt/.
7. Menze BH, Ur JA, Sherratt AG (2006) Detection of ancient settlement mounds—archaeological survey based on the SRTM terrain model. *Photogramm Eng Remote Sens* 72:321–327.
8. Altaweel M (2005) The use of ASTER satellite imagery in archaeological contexts. *Archaeol Prospect* 12:151–166.
9. Wilkinson KN, Beck AR, Philip G (2006) Satellite imagery as a resource in the prospection for archaeological sites in central syria. *Geoarchaeology* 21:735–750.
10. Deckers K, Doll M, Pfälzner P, Riehl S (2010) *Development of the Environment, Subsistence and Settlement of the City of Urkeš and its Region* (Harrassowitz, Wiesbaden).
11. Rosen AM (1986) *Cities of Clay: The Geoarchaeology of Tells* (Univ of Chicago Press, Chicago).
12. Wilkinson TJ (2003) *Archaeological Landscapes of the Near East* (Univ of Arizona Press, Tucson, AZ).
13. Akkermans PMMG, Schwartz G (2003) *The Archaeology of Syria: From Complex Hunter-Gatherers to Early Urban Societies (ca. 16,000-300 BC)* (Cambridge Univ Press, Cambridge, UK), pp 233–287.
14. Stein GJ (2004) Structural parameters and sociocultural factors in the economic organization of north Mesopotamian urbanism in the third millennium BC. *Archaeological Perspectives on Political Economies* (Univ of Utah Press, Salt Lake City), pp 61–78.
15. Ur JA (2010) Cycles of civilization in northern mesopotamia, 4400-2000 BC. *J Archaeol Res* 18:387–431.
16. Galiatsatos N (2004) Assessment of the CORONA series of satellite imagery in landscape archaeology: A case study from the Orontes valley, Syria. PhD thesis (Durham University, Durham, NC).
17. Fowler MJF (2004) Archaeology through the keyhole: The serendipity effect of aerial reconnaissance revisited. *Interdiscip Sci Rev* 29:118–134.
18. Ur JA (2003) CORONA satellite photography and ancient road networks: A northern Mesopotamian case study. *Antiquity* 77:102–115.
19. Ur JA (2010) *Urbanism and Cultural Landscapes in Northeastern Syria: The Tell Hamoukar Survey, 1999-2001* (University of Chicago Oriental Inst, Chicago).
20. Hamlin CL (1977) Machine processing of Landsat data: An introduction for anthropologists and archaeologist. *MASCA News* 13:1–11.
21. Baldridge AM, Hook SJ, Grove CI, Rivera G (2009) The ASTER spectral library version 2.0. *Remote Sens Environ* 113:711–715.
22. Breiman L (2001) Random forests. *Mach Learn J* 45:5–32.
23. Menze BH, Kelm BM, Splitthoff DN, Koethe U, Hamprecht FA (2011) On oblique random forests. *European Conference on Machine Learning and Principles and Practice of Knowledge Discovery in Databases*, , Lecture Notes in Computer Science pp:453–469.
24. Menze BH, et al. (2009) A comparison of random forest and its Gini importance with standard chemometric methods for the feature selection and classification of spectral data. *BMC Bioinf* 10:213, http://w09.biomedcentral.com/1471-2105/10/213.
25. Lahiri S (2003) *Resampling Methods for Dependent Data* (Springer, New York).
26. Ur JA, Wilkinson TJ (2008) Settlement and economic landscapes of Tell Beydar and its hinterland. *Beydar Studies I, Subartu 21* (Brepols, Turnhout, Belgium), pp 305–327.
27. Wilkinson TJ (2000) Archaeological survey of the Tell Beydar Region, Syria, 1997: A preliminary report. *Tell Beydar: Environmental and Technical Studies, Subartu VI* (Brepols, Turnhout, Belgium).
28. Wright HT, Rupley ESA, Ur JA, Oates J, Ganem E (2006) Preliminary report on the 2002 and 2003 seasons of the Tell Brak sustaining area survey. *Annales Archéologiques Arabes Syriennes* 49–50:7–21.
29. Ristvet L (2005) *Settlement, Economy, and Society in the Tell Leilan Region, Syria, 3000-1000 BC* (Univ of Cambridge, Cambridge, UK).
30. Weiss H (2003) Ninevite 5 periods and processes. *The Origins of North Mesopotamian Civilization: Ninevite 5 Chronology, Economy, Society, Subartu IX* (Brepols, Turnhout, Belgium), pp 593–624.
31. Wilkinson TJ, Tucker DJ (1995) *Settlement Development in the North Jazira, Iraq* (Aris and Phillips, Warminster, UK).
32. Ur JA, Karsgaard P, Oates J (2007) Urban development in the ancient Near East. *Science* 317:1188.
33. Ur JA, Karsgaard P, Oates J (2011) The spatial dimensions of Early Mesopotamian urbanism: The Tell Brak suburban survey, 2003–2006. *Iraq* 73:1–19.
34. Rosenstock E (2009) *Tells in Südwestasien und Südosteuropa: Untersuchungen zur Verbreitung, Entstehung und Definition eines Siedlungsphänomens* (Bernhard Albert Greiner, Remshalden, Germany).
35. Wilkinson TJ, French CAI, Matthews W, Oates J (2001) Geoarchaeology, landscape and the region. *Excavations at Tell Brak, Vol. 2: Nagar in the Third Millennium BC* (McDonald Inst for Archaeological Research, Cambridge, UK), pp 1–14.
36. Courty MA (1994) Le cadre paléographique des occupations humaines dans le bassin du Haut-Khabur (Syrie du Nord-Est): Premiers résultats. *Paléorient* 20:21–59.
37. Beck J, Sieber A (2010) Is the spatial distribution of mankind's most basic economic traits determined by climate and soil alone? *PLoS One* 5:e10416.
38. Van Liere WJ (2003) *Survey of Soil, Present Land Use and Land Capabilities of the Jezireh* (Bloomsbury, London).
39. Food and Agriculture Organization of the United Nations (1966) Étude des ressources en eaux souterraines de la Jazireh syrienne. (United Nations, Rome).
40. Hole F (2007) Agricultural sustainability in the semi-arid Near East. *Clim Past* 3:193–203.
41. Hufnagel L, Brockmann D, Geisel T (2006) The scaling laws of human travel. *Nature* 439:462–465.
42. González MC, Hidalgo CA, Barabási AL (2008) Understanding individual human mobility patterns. *Nature* 453:779–782.
43. Hodder I, Orton C (1976) *Spatial Analysis in Archaeology* (Cambridge Univ Press, Cambridge, UK).
44. Seeley J (1949) The net of reciprocal influence: A problem in treating sociometric data. *Can J Psychol* 3:234–240.
45. Sherratt AG (2010) Tellspotting. *ArchAtlas* Edition 4, Available at http://www.archatlas.org/Tellspotting/TellsMain.php. Accessed November 3, 2006.