

ARTIGO | *PAPER*

PREVENDO O PASSADO: UM PANORAMA SOBRE AS ANÁLISES ESPACIAIS, OS DADOS DIGITAIS E A MODELAGEM PREDITIVA NA ARQUEOLOGIA

PREDICTING THE PAST: AN OVERVIEW OF SPATIAL ANALYSIS, DIGITAL DATA AND PREDICTIVE MODELING IN ARCHEOLOGY

Lucas Bonald^a
Demétrio Mützenber^b
Eduardo Krempser^c

^a Universidade Federal de Pernambuco; Doutorando do Programa de Pós-Graduação em Arqueologia da UFPE; Bolsista Capes; lucas.bonald@ufpe.br

^b Universidade Federal de Pernambuco; Docente do Programa de Pós-Graduação em Arqueologia da UFPE; demetrio.mutzenberg@ufpe.br

^c Fundação Oswaldo Cruz - Fiocruz; Pesquisador em Saúde Pública – Plataforma Institucional Biodiversidade e Saúde Silvestre/Fiocruz; eduardo.krempser@fiocruz.br

RESUMO

As análises espaciais têm um destaque nos trabalhos arqueológicos, visto que são fundamentais na compreensão dos contextos e na delimitação de áreas de estudos em Arqueologia. O avanço acelerado do desenvolvimento de tecnologias da informação fez dos estudos espaciais uma peça chave para a construção do pensamento arqueológico e trouxe consigo a possibilidade de extrair uma quantidade maior de informações, seja diretamente nas atividades de campo ou em análises laboratoriais. Esse crescimento paulatino das análises espaciais traz consigo um debate a respeito do uso de meios digitais na Arqueologia, os quais precisam ser discutidos com seriedade, afinal, não é de intenção da Ciência criar “apertadores de botões”. Para isso, surge a Arqueologia Digital, com o intuito de desenvolver as bases teórico-metodológicas da Arqueologia, de modo a lidar com as aplicações computacionais. Esse ramo das ciências arqueológicas também é importante para a discussão em torno dos dados digitais e dos riscos e benesses que estes possibilitam, por isso discutiremos, aqui, como geri-los. Além disso, também será levantada a importância de termos informações abertas e irrestritas, que possibilitem uma troca maior de conhecimento entre os pesquisadores e impulsionem o avanço da Arqueologia, sobretudo no contexto digital. Com todo o progresso tecnológico, que culminou no cenário atual das análises espaciais em Arqueologia, plataformas como os SIG ganharam uma importância ímpar. O uso dessas ferramentas, aliado às legislações patrimoniais usadas para salvaguardar os vestígios arqueológicos em projetos infra estruturais, fez com que se desenvolvessem métodos preditivos a partir da busca de padrões nos dados, cujo intuito é encontrar sítios arqueológicos. Os chamados modelos preditivos arqueológicos (MPA) vêm sendo utilizados em todo o mundo e, aqui, faremos um panorama das pesquisas sobre o tema, buscando enumerar suas limitações e falar das possíveis implementações, como é o caso da Inteligência Artificial (IA), a qual tem o intuito de desenvolver modelagens que oferecem respostas mais graduadas ao arqueólogo.

PALAVRAS-CHAVE

Análise Espacial; Dados; Modelos Preditivos em Arqueologia; Arqueologia Digital.

ABSTRACT

Spatial analysis is highlighted in archaeological research, as it is fundamental in understanding the contexts and defining areas of study in Archaeology. The accelerated advance in the development of information technologies made spatial studies a key element in the construction of archaeological thinking and brought with it the possibility of extracting a greater amount of information, either directly in field activities or in laboratory analyses. This gradual growth of spatial analysis brings with it a debate regarding the use of digital media in Archaeology, which need to be discussed seriously, after all, it is not Science's intention to create "button-pushers". For this, Digital Archaeology emerges to develop the theoretical-methodological bases of Archeology to deal with computational applications. This branch of the archeological sciences is also important for the discussion around digital data and the risks and benefits that these enable, so we will discuss here how to manage them. In addition, the importance of having open and unrestricted information will also be raised, enabling a greater exchange of knowledge among researchers, and boosting the advancement of Archaeology, especially in the digital context. With all the technological progress, which culminated in the current scenario of spatial analysis in Archaeology, platforms such as GIS have gained a major importance. The use of these tools, combined with heritage legislation used to safeguard archaeological remains in infrastructure projects, led to the development of predictive methods based on the search for patterns in the data, whose aim is to find archaeological sites. The so-called archaeological predictive models (APM) have been used all over the world and here we will make an overview of researches on this subject, seeking to enumerate its limitations, talk about possible implementations, as is the case of Artificial Intelligence (AI), which has the aim of developing models that offer more reliable answers to the archaeologists.

KEYWORDS

Spatial analysis; Data; Archaeological Predictive Model; Digital Archaeology.

COMO CITAR ESTE ARTIGO

BONALD, Lucas; MÜTZENBERG, Demétrio; KREMPSE, Eduardo. Prevendo o passado: um panorama sobre as análises espaciais, os dados digitais e a modelagem preditiva na arqueologia. Cadernos do Lepaarq, v. XIX, n.38, p. 40-63, Jul-Dez. 2022.

INTRODUÇÃO

A busca por compreender as relações entre os grupos humanos do passado e o meio ambiente em que estes estavam inseridos é um debate que há muito tempo ocorre na Arqueologia. Desde o período dos colecionistas, no século XVI, a noção de escala e análises espaciais em níveis locais foi concebida em atividades de campo. A cartografia também foi inserida às primeiras pesquisas arqueológicas aplicadas em trabalhos na Grã Bretanha, como auxílio na localização de sítios arqueológicos (GILLINGS; HACIGUZELLER; LOCK, 2019).

A partir do século XIX, porém, o meio ambiente passa a ser visto como uma variável essencial para entender como viviam os grupos pretéritos. Thomsen e Worsaae, por exemplo, já incluíam os contextos ambientais nos estudos arqueológicos da península escandinava (TRIGGER, 2011), mas somente no século XX, com os trabalhos de Grahame Clark (1952) e Steward (1955), o enfoque ambiental na Arqueologia ganha bases teóricas mais sólidas. Clark atrelava a mudança cultural a questões de instabilidade temporal, devido a mudanças ambientais, nas taxas populacionais, nas inovações e no contato. Já Steward, que trabalhava na perspectiva de uma “evolução cultural”, observou que as variações culturais só seriam perceptíveis por meio de uma análise a respeito da adaptação do ser humano ao meio.

Os estudos espaciais em Arqueologia, desde seus primórdios, carregam uma base teórico-metodológica que advém das Geociências, com isso, alguns conceitos-chaves servem de base para a construção acerca do debate arqueológico sobre o espaço. Para exemplificar, partiremos do que a Sociedade Real de Geografia chama de “conceitos de alto nível” (GREGORY; LEWIN, 2018), dos quais o “espaço” é um deles.

Dentro da Geografia Humana, Santos (2008) define o espaço como um local onde homem e meio se relacionam de forma dinâmica, ou seja, é um sistema de valores que está em constante transformação no que diz respeito à função e ao significado dos seus objetos. Em complemento à definição anterior, é possível afirmar que o espaço é, antes de tudo, uma estrutura social que guarda aspectos intrínsecos da sociedade, como as relações de poder, os sistemas políticos e todos os tipos de inter-relações entre o homem e o meio (KOOOPS; GALIČ, 2017).

Ao olharmos para a Arqueologia, o conceito do espaço é permeado por cinco aspectos primordiais, sendo eles: é ligado ao tempo; é difícil de ser representado; está relacionado com as práticas diárias de um grupo; está atrelado à mobilidade; e trata muito mais de ausências do que de presenças (GILLINGS; HACIGUZELLER; LOCK, 2019).

Destacando sobretudo os dois últimos pontos, podemos dizer que as ausências acima relatadas partem do ponto de que os vestígios arqueológicos e o espaço onde estes estão inseridos não se limitam apenas à área definida convencionalmente como “sítio arqueológico”, pois, locais em que não há vestígios, mas ficam em áreas nas periferias de concentrações, certamente tratam-se de ambientes em que os grupos humanos do passado estiveram. Afinal, o espaço é, antes de tudo, atrelado à mobilidade, já que essas estruturas sociais precisavam realizar suas atividades de subsistência, o que demandava uma migração periódica (SCHIFFER, 1996; VERHAGEN;

WHITLEY, 2011).

Desde a década de 1970, Plog, Plog e Wait (1978) já enxergavam essa noção não limitante de sítios arqueológicos. Eles denominaram de “não-sítio” os locais onde há presença de atividade humana, porém não é possível delimitá-los espacialmente. Essa visão foi corroborada pelos estudos de Schiffer (SCHIFFER, 1996) a respeito da Formação do registro arqueológico, quando é lançado luz sobre os artefatos descartados e o reuso, indicando essa dinamicidade no registro arqueológico e se opondo à visão materialista e estática de Willey e Phillip (1958) e Binford (1964), que viam a necessidade de haver uma concentração vestigial para denominar o espaço de sítio arqueológico.

Tendo o sítio arqueológico como peça fundamental, Vita-Finzi e colaboradores (1970), desenvolveram a chamada análise de captação de sítios, em que buscava-se compreender a respeito da exploração de recursos por parte dos grupos humanos antigos, utilizando análises espaciais em grandes zonas que partiam de um centro específico. Mesmo anterior às discussões sobre esse conceito de sítio mais abrangente, os autores já tinham uma visão dinâmica e ampla do espaço, trazendo as necessidades humanas à discussão. É a partir desse tipo de análise que surgem os primeiros modelos preditivos em Arqueologia, procurando por padrões ambientais que se relacionavam aos sítios arqueológicos.

Análises como as de captação de sítios e outros tipos de estudos estatísticos e espaciais na Arqueologia têm ganhado espaço com o passar dos anos e os avanços tecnológicos. Os trabalhos em campo e as análises laboratoriais têm gerado uma quantidade de dados muito grande (HUGGETT, 2020). No que diz respeito aos estudos do espaço especificamente, o cenário não é diferente, já que tivemos melhorias na topografia, no sensoriamento remoto, na fotogrametria e na geofísica (DOUGLASS; LIN; CHODORONEK, 2015; HUGGETT, 2017).

Para lidar com esses dados espaciais, frutos dos avanços tecnológicos, as ciências que analisam o espaço têm recorrido aos Sistemas de Informação Geográfica (SIG). Segundo Câmara et al (1996), o SIG é um sistema capaz de analisar, manipular e armazenar dados geográficos. Em complemento, na Arqueologia, este é definido como “uma plataforma para obter e integrar arquivos de dados espaciais, como mapas topográficos, localização e morfologia dos sítios, fotografia aéreas (...) tudo isso integrado a um meio analítico comum.” (CONOLLY; LAKE, 2009, p. 31).

A inclusão dessas novas técnicas e tecnologias ao cotidiano arqueológico não deve, porém, limitar a capacidade criativa dos pesquisadores, que precisam fazer de suas metodologias, aportes inovadores e não meras replicações advindas de outras áreas do conhecimento. É preciso, então, desenvolver uma base teórico-metodológica específica para Arqueologia, que se utilize das ferramentas, como os SIG, para propor novas ideias de ação, as quais possam ser aproveitadas inclusive por outras áreas do conhecimento (VERHAGEN, 2018).

Nestes novos tempos em que a Arqueologia trabalha com muitas informações e necessita do auxílio de ferramentas computacionais para a realização de uma gestão de dados mais cuidadosa, torna-se ainda mais necessário que as discussões em torno das técnicas e tecnologias utilizadas sejam inseridas ao debate acadêmico, seja em sala de aula ou congressos, contribuindo

do, assim, para eliminar a realidade dos “apertadores de botões”, os quais desenvolvem suas metodologias de forma exclusivamente operacional, e não entendem os pormenores básicos do método usado (HUGGETT, 2017; HUGGETT; REILLY; LOCK, 2018; ZUBROW, 2006).

Conhecer profundamente o arcabouço ferramental aplicado evita não só a simples replicação de rotinas, mas também fornece uma base para repensar a Arqueologia. É o que coloca Huggett (2017), quando define os chamados “artefatos cognitivos”, que podem ser desde uma colher de arqueólogo até uma estação total ou um computador, e têm a função de complementar a cognição arqueológica, possibilitando ao pesquisador enxergar elementos que seriam impossíveis apenas com a capacidade humana.

Por fim, vale ressaltar que os artefatos cognitivos, sobretudo os equipamentos eletrônicos, tais como máquinas fotográficas; computadores; drones; receptores GNSS; entre outros, geram dados digitais, que podem trazer ganhos substanciais às pesquisas arqueológicas. Por outro lado, carregam consigo o risco de uma supervalorização da técnica, o que só pode ser combatido com uma base teórica consolidada (HUGGETT, 2017).

ARQUEOLOGIA DIGITAL

Os estudos espaciais se desenvolveram graças às ferramentas como os SIG, porém, para construir tal plataforma, são necessários *softwares* capazes de ter em sua interface um conjunto de funções propícias para análise espacial, juntamente a um *hardware* robusto por trás, capaz de processar todas as atividades de cálculos desejados. Portanto, os avanços tecnológicos, especialmente na área da computação, permitiram o cenário em que vivemos atualmente, fornecendo os subsídios para uma compreensão e representação melhor do espaço arqueológico (FITZ, 2008; WHEATLEY; GILLINGS, 2002).

A união da Arqueologia e dos meios virtuais torna necessário o desenvolvimento de bases teóricas mais sólidas (ZUBROW, 2006). A chamada Arqueologia Digital surge dessa necessidade de estreitar as relações entre uma disciplina de base social com a exatidão da computação. Segundo Torres (2017), esse ramo da Arqueologia pode ser definido como a capacidade de diferentes técnicas e/ou tecnologias de serem aplicadas às diversas etapas dos trabalhos arqueológicos.

A Arqueologia Digital tem um apelo técnico bem característico, por ser associada diretamente à tecnologia, a equipamentos e ao maquinário, em geral, que auxiliam o arqueólogo no registro de campo. Segundo Huggett (2015), a Arqueologia Digital tem se mostrado cada vez menos como apenas um apanhado de conhecimentos emprestados de outras ciências e mais como capaz de desenvolver novas técnicas e ferramentas que sirvam não só para uso na Arqueologia, mas também em outros campos científicos. Isso indica que o cenário atual, apesar de ser inicial — sobretudo no Brasil — caminha no sentido de fornecer subsídios para as resoluções de diversos problemas metodológicos.

Ao olharmos a Arqueologia Digital de um ponto de vista mais específico, podemos dividi-la em duas vertentes, sendo a primeira voltada para estudos que envolvem um sistema de dados es-

táticos e não modificáveis, denominada de Arqueologia Virtual, e a segunda é a Cyber-Arqueologia, justamente o oposto da anterior, trabalhando com dados dinâmicos, capazes de gerar ambientes virtuais que simulam a realidade. Em outras palavras, podemos dizer, respectivamente, que, em uma são feitas reconstruções estáticas e na outra simulações do passado. (OLIVITO; TACCOLA; ALBERTINI, 2016).

As análises espaciais tratadas nos SIG lidam com os chamados dados estáticos e não modificáveis, o que as faz ser parte integrante da chamada Arqueologia Virtual. A plataforma recebe primordialmente dois tipos distintos de informações, os vetores, que se estruturam em pontos, linhas e polígonos, obtidos mediante uma vetorização ou um redesenho de determinado arquivo. Por outro lado, os outros arquivos suportados são chamados de matriciais ou *raster*, formados por *pixels*, pequenas matrizes de cores diferentes que dão forma a uma imagem (FITZ, 2008; WHEATLEY; GILLINGS, 2002).

Os estudos que utilizam os SIG para análises espaciais são frutos de muita crítica devido ao determinismo ambiental atrelado ao seu uso, em que os elementos que compõem o estudo arqueológico são representados por simples vetores ou imagens, dando a eles uma visão bastante positivista. Ultimamente, há um esforço em integrar aspectos de cunho social e cultural às plataformas (HUGGETT, 2015; SUPERNANT, 2017), como as análises de visibilidade (GILLINGS, 2015) e de acessibilidade (VERHAGEN et al., 2012).

Comentando um pouco a respeito das duas análises acima, temos que o índice de visibilidade é calculado a partir de um modelo digital de elevação e tem o intuito de mostrar o campo visual desde uma determinada localização dentro de um dado terreno. A acessibilidade, por sua vez, tem o intuito de compreender as movimentações dos povos pretéritos por meio de análise espacial baseada em caminhos de menor custo espalhados ao longo do espaço, os quais geram rotas de acesso a um ponto de interesse específico (GILLINGS, 2015; VERHAGEN, 2010).

O mundo digital aplicado às análises espaciais em Arqueologia tem, para além do determinismo ambiental, alguns pontos que precisam receber especial atenção ao se fazer Arqueologia Digital. Alguns desses já foram apresentados aqui, como a necessidade de se desenvolver uma base teórica ou novas técnicas, sem que haja uma reprodução trazida de outras ciências. Para além desses, Huggett (2015) afirma que para realizar uma grande mudança nos estudos das aplicações computacionais nas Ciências Arqueológicas é necessário, primeiramente, que os pesquisadores queiram utilizar os recursos gerados, ou seja, é preciso que eles despertem interesse para tal e para isso precisam ter uma linguagem acessível e uma estrutura bem definida. Por fim, é fundamental que haja um engajamento de toda uma comunidade arqueológica em torno de uma estrutura cooperativa para o uso dessas aplicações.

O desenvolvimento da Arqueologia Digital certamente recairá na Ciência Arqueológica como um todo, pois ela é parte integrante deste período de evolução tecnológica e, saber lidar com estes elementos do mundo moderno, faz com que algumas armadilhas não sejam disparadas, afinal, a estruturação de bases teóricas fortes aliadas a um entendimento de pormenores técnicos, nos faz melhor compreender o mundo que nos cerca, seja ele presente ou mirado a um passado.

A GESTÃO DOS DADOS DIGITAIS

Como temos falado ao longo deste artigo, os avanços tecnológicos trouxeram uma gama de possibilidades na obtenção dos dados de uma pesquisa arqueológica. Essa revolução vem levantando debates acerca da chamada “era da Big Data”. Em face disso, Huggett (2020) fala a respeito dos novos paradigmas que os estudos arqueológicos vêm passando devido a esse crescente ganho de informações. Para o autor, a virtualização dos dados traz ganhos substanciais às pesquisas arqueológicas, pois torna as informações mais flexíveis, pesquisáveis, compartilháveis, calculáveis, tudo de uma forma mais rápida.

Por outro lado, há alguns cuidados que precisam ser levados em conta quando se lida com uma grande quantidade de dados digitais. A compreensão de que os dados são simplesmente meios para o fim é um deles; além deste, as reduções e compressões de arquivos podem gerar perdas de informações, portanto, saber lidar com isso é fundamental; alguns conhecimentos extras de computação podem ser exigidos; por fim, o uso deliberado das técnicas e tecnologias pode ser um problema operacional grave, pois é preciso haver um propósito claro para a utilização de alguns equipamentos e algumas metodologias nos trabalhos arqueológicos, caso contrário, configurar-se-ia como uso pelo uso (HUGGETT, 2020).

Um outro grande problema ocasionado pelos dados é a sua manipulação deliberada com a intenção de mascarar uma realidade, ou seja, enviesá-los a bel prazer. No clássico livro escrito por Darrel Huff, “Como mentir com estatística” (HUFF, 1956), são levantados diversos exemplos de como é possível utilizar informações de forma tendenciosa, seja na Política, na Medicina e nas Ciências Humanas, como forma de ludibriar a sociedade ou ter acesso a determinados privilégios.

Por mais que os exemplos do livro citado não sejam diretamente ligados à Arqueologia, eles servem como forma de alerta para mostrar os riscos de um inabilitado trabalho com a estatística. Além disso, os modelos estatísticos, em qualquer área, são manipulados por um ser humano, que tem ideias próprias, que muitas vezes pode passar tanto para a composição do modelo, como para suas análises. Cabe, portanto, a cada pesquisador, perceber quanta interferência “externa” está sendo incluída em suas interpretações (GLOCK, 1995).

A manipulação errônea dos dados também pode causar danos sociais irreparáveis. As grandes empresas de tecnologias têm usado dos nossos dados para comercializá-los, de modo que propagandas nos atinjam com o intuito de consumirmos determinado produto. Segundo Morozov (2018), essa lógica tem feito as pessoas se tornarem reféns da tecnologia, pois essas práticas vêm moldando até nosso modo de vida. Trazendo esse debate para Arqueologia, isso é tão perigoso quanto, pois o Patrimônio Cultural está sob ameaça, caso algumas técnicas, tais como as modelagens matemáticas em Arqueologia, sejam usadas de maneira incoerente, afinal, estamos lidando com um bem durável, mas que muitas vezes são achados em estados de conservação que demandam certos cuidados.

DEMOCRATIZANDO OS DADOS

Um outro fator que precisa ser discutido com relação a todo esse crescente acúmulo de informações no cenário dos dados digitais e no aprimoramento de técnicas e metodologias, nas ciências de um modo geral, é a “democratização dos dados”, ou seja, o ato de tornar as bases de dados acessíveis a todos, de maneira gratuita. A falta de acesso às informações são um dos grandes problemas dentro da Arqueologia, e isso dificulta, inclusive, a troca de informação entre os pesquisadores. Portanto, para que haja um desenvolvimento maior nas pesquisas arqueológicas, se faz necessário que os dados das pesquisas se tornem abertos (MARCHETTI et al., 2018; MARWICK et al., 2017).

A divulgação desse tipo de informação não passa apenas pelos dados brutos, os quais servem para construir a base estatística ou uma análise espacial dos trabalhos, mas também os próprios artigos, que ainda hoje são ligados a revistas que têm uma plataforma paga, as quais exigem do autor uma taxa para tornar uma publicação aberta ou o leitor precisa desembolsar uma quantia para poder ler tal obra (MARWICK et al., 2017). Para isso, plataformas como o *Research Gate* e o *Academia.edu* têm se tornado grandes aliadas na divulgação de trabalhos científicos de forma gratuita (TORRES, 2017).

O acesso aberto aos dados é de fundamental importância quando se fala em modelagens matemáticas aplicadas à Arqueologia, já que estas trabalham com uma grande quantidade de informações que compõe as variáveis a serem usadas nos cálculos estatísticos para obtenção dos produtos. Para a construção de Modelos Preditivos Arqueológicos, por exemplo, a presença de plataformas SIG abertas como o QGIS (QGIS, 2021), integradas a gerenciadores de bancos de dados, como o PostgreSQL (POSTGRESQL, 2021), facilitam a aglutinação dos dados espaciais com dados estatísticos.

A popularização de linguagens de programação de código aberto, tais como o Python e R (PYTHON.ORG, 2018; RFOUNDATION, 2021), tem proporcionado uma evolução nas análises espaciais e estatísticas em Arqueologia, e representam mais de 30% no mercado global da tecnologia da informação, segundo o Popularity of Programming Language Index (PYPL, 2021), que é medido por meio das buscas pelos nomes das linguagens na plataforma Google. Boa parte da aplicação dessas linguagens está atrelada à Ciência dos Dados, que como o nome já sugere, se preocupa com análises e o entendimento de bases de dados, relacionadas a fenômenos humanos, sociais e naturais (HAYASHI, 1998).

Essas duas linguagens de programação supracitadas têm bibliotecas especializadas para auxiliar os utilizadores a criarem as estatísticas de seus modelos. O Python, por exemplo, tem inúmeras bibliotecas que podem ser usadas para a análise de dados e/ou desenvolvimento de modelos de aprendizagem de máquina, tais como o *Scikitlearn*, *Tensorflow*, *Keras* e *PyTorch*, além dessas, há outras auxiliares, como o *Pandas*, o *Matplotlib* e o *Numpy* (GÉRON, 2019).

Dentro da linguagem R, que é um projeto especializado para o desenvolvimento de modelos estatísticos e na análise dos dados, as bibliotecas acabam sendo muito diversas, pois existem,

muitas vezes, pacotes bem específicos, para cada determinado tipo de algoritmo, porém, um dos mais usados são o *CARET*, *alrr3*, *lmtest*, *arules*. Além desses, um grande suporte, a qualquer análise de dados, é a geração de gráficos, que no R, tem na biblioteca *ggplot2*, seu grande trunfo (LESMEISTER, 2015).

Entre essas duas linguagens trazidas, o R ainda tem bastante destaque dentro das Ciências Humanas e, conseqüentemente, na Arqueologia. Segundo uma pesquisa feita por Schmidt e Marwick (2020), existem diversas revistas de Arqueologia nas quais é possível ver a citação a tal linguagem, e ao longo dos anos, a inserção desse assunto vem sendo ampliada. Um outro ponto observado pelos autores é que boa parte dos artigos que usam do R estão ligados a estudos que envolvem uma análise ambiental, o que faz sentido, justamente pelo fato desta possuir mecanismos e bibliotecas voltadas para o uso de informações geográficas, bem como a possibilidade de integração com o QGIS.

Ainda no âmbito da discussão a respeito da abertura de dados, as linguagens de programação e *sites* como o Github (GITHUB, 2021) favorecem a divulgação e troca de informações por meio de códigos e/ou modelos aplicados à Arqueologia (SCHMIDT; MARWICK, 2020), mas estes ainda precisam ser mais utilizados, pois ainda hoje a quantidade de pesquisadores de Arqueologia que usam essa plataforma é muito pequena.

Compreender os riscos atrelados aos dados digitais, aliados a um acesso aberto e compartilhado de dados das diversas pesquisas realizadas na Arqueologia, dão subsídios ao desenvolvimento de grandes redes de informação, que poderão ser usadas na difusão de metodologias para melhor compreender a relação entre os grupos humanos pretéritos e o meio ambiente. Uma das principais são as modelagens preditivas em Arqueologia (MPA), que têm ganhado espaço devido à expansão de implantações infraestruturais e conseqüente regulação por legislações ambientais (KIPNIS, 1997; VERHAGEN, 2018).

OS MODELOS PREDITIVOS EM ARQUEOLOGIA (MPA)

A história do uso de modelos preditivos, de um modo geral, remonta ao início da era dos computadores, por volta da década de 1940, do século XX. À medida que essas máquinas iam evoluindo em suas capacidades de processamento tornou-se possível a aplicação de modelos computacionais mais complexos. Com isso, órgãos governamentais são os primeiros a se utilizar de análises de dados como forma de prever algo (VAN RIJMENAMM, 2013).

Como os computadores têm avançado constantemente desde o período acima citado, a discussão em torno de uma definição a respeito da modelagem preditiva está em contínuo desenvolvimento. Kunh e Johnson (2013) a conceituam como “o processo de desenvolvimento de uma ferramenta ou modelo matemático que gera uma predição acurada.”

A modelagem preditiva é uma das utilizações computacionais ou estatísticas mais antigas dentro da Arqueologia, muito desse pioneirismo se deve sobretudo ao fato de a Arqueologia ser, antes de tudo, uma Ciência Espacial, como apontado por Gillings, Haciguzeller e Lock (2019). Se-

gundo Verhagen (2018), os Modelos Preditivos em Arqueologia (MPA) são meios de prever onde estão materiais e sítios arqueológicos, baseados em amostras de dados já conhecidas ou a partir do entendimento de padrões de comportamento humano.

Na definição acima, a necessidade de saber como os vestígios arqueológicos estão dispostos no ambiente fica evidente, mas o final é que chama a atenção, pois, também, não se deve esquecer que além da preocupação espacial, o contexto social também é necessário quando se pensa em modelar o ser humano, já que este tem um comportamento particular. Como forma de inserir essas variáveis ao MPA, podemos recorrer aos estudos de Verhagen (2012), acerca da acessibilidade, e de Gillings (2015), a respeito da visibilidade, anteriormente levantados nesta discussão.

Buscando justamente um desenvolvimento de um modelo como solução, pelo menos parcial, para esse problema do determinismo ambiental, Verhagen e Whitley (2011) propõem o uso do que eles chamam de Modelo Preditivo Cognitivo, o qual se baseia em conceitos da Arqueologia Cognitiva (RENFREW; ZUBROW, 1994) e na Arqueologia Comportamental (SCHIFFER, 2010), buscando colocar o Homem como um ser que tem suas escolhas e vontades, ou seja, trazer os elementos individuais do comportamento humano, aliado a uma percepção do espaço que busca observar o ambiente de uma forma ampla e o registro arqueológico como elemento que pode ser submetido a agentes sincrônicos e pós deposicionais.

Os processos cognitivos trazidos por Verhagen e Whitley (2011) se baseiam na questão de tomadas de decisão, que podem, por sua vez, ser divididos em dois grupos de variáveis, as diretas, as quais podem ser percebidas com os sentidos humanos, e as indiretas, representadas pelas informações adquiridas por meio da experiência e especulação.

Os primeiros MPA, atrelados ainda a uma visão descritiva e fortemente influenciados pelas variáveis ambientais, foram desenvolvidos em uma pesquisa no Peru, por Gordon Willey (1953), que buscava compreender a distribuição espacial dos assentamentos pré-históricos do Vale do Virú. É com a ideia de análise de captação de sítios, por meio das pesquisas de Vita-Finzi et al (1970), que percebe-se que é possível se extrair padrões através das análises que correlacionam os dados arqueológicos e ambientais e, a partir disso, criar previsões. Em paralelo às pesquisas de Vita-Finzi e seus colaboradores, os MPA começam a se popularizar no contexto da CRM (*Cultural Resource Management*), que compreende os trabalhos de Arqueologia de Contrato do cenário estadunidense. O crescimento acelerado do país da América do Norte, aliado a leis mais rígidas na defesa do patrimônio arqueológico, fez dos MPA um grande auxiliar (VERHAGEN, 2008; VERHAGEN; BORSBOOM, 2009; BARCELÓ; DEL CASTILLO, 2016).

Somando ao contexto infraestrutural, a Arqueologia também desenvolvia bases teóricas alinhadas a um pensamento hipotético-dedutivo, buscando trazer mais empirismo aos trabalhos, e em consonância à Teoria dos Sistemas. Essas características típicas da Arqueologia Processual faz as pesquisas se tornarem mais interdisciplinares e aspectos trazidos de ciências como a Geografia, a Matemática e a Estatística tornam-se mais comuns dentro das Ciências Arqueológicas (CLARKE, 1968; ROBRAHN-GONZÁLEZ, 2000).

Uma das primeiras obras a tratar da modelagem na Arqueologia é o livro *Models in Archaeology*, editado por David Clarke. A coletânea de artigos visava trazer de forma mais explícita como os modelos poderiam auxiliar na compreensão dos grupos humanos do passado (CLARKE, 1972). A figura de Clarke é importante por trazer para discussão arqueológica todas essas pautas, e nesse período cada vez mais interdisciplinar, uma das principais ciências de qual a Arqueologia se aproximou foi a Estatística, a qual já vinha sendo usada em trabalhos anteriores (MYERS, 1950; SPAULDING, 1953), mas é apenas durante as décadas de 1960 e 1970 que ela ganha bases mais sólidas, permitindo relacionar de forma mais empírica algumas das informações ambientais com os locais referentes ao sítios arqueológicos.

Na modelagem preditiva, os dados podem ser interpretados de duas formas, caracterizando assim dois tipos de modelos. Segundo Verhagen e Whitley (2020), um se baseia na dedução, ou seja, a partir de alguns pressupostos teóricos ou em informações preconcebidas ou preestabelecidas, como por exemplo, a necessidade de se hidratar ou se alimentar, de um ser humano, serviria de base para buscar sítios arqueológicos próximos a áreas com fontes de água. Já o segundo tipo, se baseia na indução, ou seja, em fatos já conhecidos e testados, guiados, muitas vezes, por implementações estatísticas. Um exemplo desses modelos podem ser quando se parte em busca novos sítios, como base na existência de outros já conhecidos.

Dentro dos modelos indutivos, podemos destacar ainda dois tipos de composições diferentes. Esses podem ser compostos apenas de uma única classe, que representam os sítios arqueológicos já conhecidos e, a partir da localização espacial dos elementos que a integram, é calculada a relação estatística destes com as variáveis ambientais ou sociais associadas, de modo a encontrar os padrões. Por outro lado, os modelos indutivos podem ser formados por duas classes ou mais. Além dos elementos que se referem aos sítios arqueológicos, haverá também áreas de ausência de vestígio, o que fará o modelo trabalhar na dicotomia ausência/presença (VERHAGEN; WHITLEY, 2020).

Adentrando mais nesse segundo espectro de modelos indutivos, precisamos retornar ao debate acerca do conceito de sítio arqueológico, afinal, será necessário usá-lo aqui com o propósito de bem delimitar as áreas de presença e ausências. Além disso, o fato de se trabalhar com ausência levanta outra problemática associada aos trabalhos de prospecção arqueológica. É preciso incluir nas metodologias que não só os dados com vestígios arqueológicos são de interesse da pesquisa, mas não ter “nada” também é de suma importância para a construção de modelagens mais precisas.

Um dos grandes exemplos dessa aplicação é o MnModel (HOBBS, 2019), o qual é fruto de um projeto ligado ao Departamento de Transporte do Estado de Minnesota, nos EUA. No ano de 2019, a fase mais recente (Fase 4) do programa foi publicada, indicando uma eficiência de cerca de 90% no encontro de diversos sítios arqueológicos da região. Esses resultados só foram alcançados depois de mais de 20 anos de desenvolvimento do MPA, que foi ampliando sua base de dados e as escolhas nas técnicas utilizadas. Segundo apuraram Verhagen e Whitley (2020), nos últimos 7 anos, o projeto custou 4 milhões de dólares aos cofres públicos, mas conseguiu economizar 3 milhões por ano, o que mostra o quão vantajoso é investir em projetos desse tipo (HOBBS, 2019).

A APRENDIZAGEM DE MÁQUINA E OS MPA

Um modelo como o MnModel passou por diversas etapas ao longo dos anos e diferentes métodos estatísticos foram testados para a geração de produtos. Na sua última fase, os pesquisadores utilizaram uma metodologia baseada no uso de algoritmos¹ de Inteligência Artificial (IA), mais especificamente a Aprendizagem de Máquina (AM) ou *Machine Learning*, por meio do emprego de Florestas Aleatórias ou *Random Forests* (HOBBS, 2019).

Para não haver nenhum mal entendido em achar que as máquinas irão substituir os arqueólogos, aqui é preciso esclarecer alguns pontos. Primeiramente, a IA pode ser entendida como uma técnica que proporciona uma máquina a reproduzir a inteligência humana (CERON, 2019). Já quando falamos em AM, que é um ramo da IA, temos, então, um processo automatizado em que se extrai padrões por meio dos dados (KELLEHER; MACNAMEE; D'ARCY, 2015). Essa prática aos poucos vem ganhando espaço nos estudos arqueológicos, sobretudo pelo fato de as principais linguagens computacionais oferecerem bibliotecas com suporte direto a elas.

Adentrando ainda mais no espectro da AM, temos alguns tipos de “aprendizados”. São eles os supervisionados, não supervisionados, os semi-supervisionados e o por reforço. Aqui neste artigo vamos tratar dos dois primeiros, afinal, estes estão presentes na grande maioria dos trabalhos que envolvem modelagem preditiva em diversas áreas do conhecimento (GÉRON, 2019).

O aprendizado supervisionado é caracterizado quando dados alimentados pelo algoritmo contêm as informações ou os rótulos de toda a base de dados previamente levantada e usa dessas informações para aprender e obter respostas para os dados não rotulados ou não conhecidos (BURKOV, 2019). Trazendo isso para um exemplo arqueológico, podemos dizer que quando temos os dados de sítios arqueológicos e usamos destes para aprender sobre as informações fornecidas e com isso gerar mais dados que ainda não eram conhecidos, temos uma aprendizagem supervisionada.

Podemos dividir a aprendizagem supervisionada em duas categorias de algoritmos, a classificação e a regressão, sendo a primeira composta de rótulos de classes finitas, associados geralmente a algum tipo de característica nominal, podendo apresentar duas ou mais classes, muito usada, por exemplo, na distinção entre e-mails “spams” ou não. Por outro lado, a regressão trata da previsão de valores reais e contínuos, sobretudo números, tendo como exemplo comum ao cotidiano a previsão no preço de um imóvel, levando em conta as características físicas deste, como tamanho, localização e número de aposentos (BISHOP, 2011; BURKOV, 2019; GÉRON, 2019).

O aprendizado supervisionado é destaque nos modelos preditivos arqueológicos usados atualmente, sendo também o mais comum. Dentre os principais tipos existentes hoje, temos: a Floresta Aleatória — usada do MnModel —, a regressão logística, as redes neurais artificiais e as máquinas de vetores suporte (GÉRON, 2019).

O outro tipo de aprendizado destacado, o não supervisionado, como nome já sugere, não

¹ São uma sequência de etapas computacionais que transformam os dados de entrada em dados de saída ou ainda uma ferramenta para resolver problemas computacionais específicos (CORMEN et al., 2001)

se utiliza das informações já conhecidas e aprende apenas com as variáveis dadas (ABU-MOSTAFA; MAGDON-ESMAIL; LIN, 2012). Esse tipo de aprendizado vem sendo usado na Arqueologia como forma de criar categorias tipológicas em estudos de escala mais micro, como no caso dos grafismos rupestres ou das cerâmicas (ANICHINI; WRIGHT; VILA, 2018; BANERJEE; SRIVASTAVA; PIKE, 2018; PUGIN, 2016).

Os algoritmos de AM, que segundo Barceló (2009) aparecem na Arqueologia com a capacidade de fazer as máquinas aprenderem com inferências e erros postos a elas, têm a capacidade de fornecer respostas mais graduadas ao arqueólogo, retirando parte do viés, já que o mesmo não exige que se definam pesos às variáveis — calculados durante processo — e além do mais, fornecem técnicas de seleção de atributos, fazendo com que os modelos possam ser trabalhados com diversas variáveis, valendo-se da aplicação de “filtros” que irão identificar os elementos estatisticamente mais significativos.

A aplicação dessas ferramentas não são a solução para todos os problemas relacionados aos MPA, e por se tratar de uma modelagem, estes são apenas formas de melhorar o poder explicativo dos eventos históricos, sem que isso seja uma verdade absoluta (BARCELO, 2018). Além disso, há outras críticas e outros problemas a serem resolvidos nos MPA.

Um destes problemas é mais geral e já vem sendo criticado há muito tempo. Os MPA têm seus resultados expressos em probabilidades, ou seja, eles indicam locais com alta, média e baixa probabilidade de se achar um sítio arqueológico. Isso tem feito com que pesquisadores usem a modelagem para procurar sítios apenas em áreas de alta probabilidade e esqueçam totalmente dos outros locais, o que é contraindicado, pois estamos tratando de previsões de um passado, a partir de algumas variáveis que jamais poderão representar com fidedignidade o comportamento humano pretérito (VERHAGEN; WHITLEY, 2011).

Essa prática pode levar a perdas irre recuperáveis do patrimônio arqueológico, por isso, por mais que tenhamos nossos resultados digitais, para que estes sejam atestados, é preciso haver idas ao campo que possam não só inferir a respeito da qualidade do modelo construído, mas também no intuito de levantar mais informações e assim ampliar a base de dados. Modelagens preditivas são construções que levam anos a serem aprimoradas e não podemos acreditar cegamente nos dados digitais, pois estaríamos caindo na armadilha levantada por Morozov (2018), no tópico anterior deste trabalho, quando o autor fala acerca do condicionamento social criado pelas grandes empresas de tecnologias.

O outro problema, esse mais atrelado aos MPA que usam da AM e mais especificamente os que trabalham com a dicotomia ausência/presença, é que ao analisarmos friamente o que estes nomes representam na Arqueologia, temos de cara uma dificuldade. Afinal, a ausência representa locais visitados, mas que não há sítios, já a presença, configura-se o oposto disso, e está atrelada diretamente aos sítios arqueológicos. Em um mundo real, porém, haverá além dessas duas classes, uma “massa cinzenta” de informações que não se conhece a respeito, e é sobre ela que serão observadas as previsões. Com isso em mente, de todo o modo, temos a certeza de que

teremos muito mais áreas sem sítios arqueológicos do que o contrário.

O fato de existir uma classe que aparece muito mais que outra é um grande problema para a AM no geral, não é à toa que a Ciência da Computação vem desenvolvendo formas de lidar com isso. Eles denominam esse tipo de configuração de um *dataset*, como uma “base de dados desbalanceada”, a qual pode ser amenizada por meio do uso de algumas implementações estatísticas (FERNÁNDEZ et al., 2018).

A principal implementação estatística a ser usada para lidar com dados desbalanceados são os métodos de reamostragem. Podemos destacar aqui dois tipos distintos destes, um chamado de *oversampling* ou sobreamostragem, caracterizado por um balanceamento a partir da classe minoritária, buscando igualá-la à majoritária. O outro, por sua vez, chamado de *undersampling* ou subamostragem, tem o objetivo de atuar na classe majoritária com o intuito de equipará-la à minoritária (BROWNLEE, 2020).

Diversos são os exemplos que podemos relatar na Arqueologia que usam dos tipos de aprendizados citados aqui. A classificação tipológica de vestígios arqueológicos é o ramo mais explorado, no quesito de uso automatizado de processos nas Ciências Arqueológicas (BARCELÓ, 2010) usando para isso o aprendizado não supervisionado. Um grande exemplo de aplicações nesse ramo classificatório-tipológico é o projeto ArchAIDE, que visa identificar fragmentos cerâmicos a partir de uma foto (ANICHINI; WRIGHT; VILA, 2018). No ramo das análises do material lítico, Flores et al (2019) também desenvolveram um sistema de classificação, porém para pontas de projéteis.

Dentro do estudo de práticas para o reconhecimento de novos sítios arqueológicos e na aplicação de MPA, são inúmeros os trabalhos que se utilizam da AM, como Caspari e Crespo (2019), Orengo e Arnau (2019), Kramer (2020), Berganzo-Besga et al (2021), que buscam identificar, de maneira automatizada, em imagens de satélites, marcas nos solos que indiquem a presença de sítios arqueológicos. Há também pesquisas no sentido de criar modelagens utilizando aspectos ambientais extraídos de imagens *raster* e transformadas em números, como forma de servir de base para cálculo e o encontro de novos sítios arqueológicos (ANNA NILSSON, 2016; KLEHM et al., 2019; MALAPERDAS; ZACHARIAS, 2019; NOVIELLO et al., 2018; WACHTEL et al., 2018; BANERJEE; SRIVASTAVA, P.; PIKE, 2018; PUGIN, 2016).

O CENÁRIO DOS MODELOS PREDITIVOS NO BRASIL

No Brasil, o cenário para a implementação de modelos preditivos começa a ser discutido justamente no período de grande aumento nas pesquisas para implantação de obras de engenharia civil de grande impacto. Um artigo seminal para a discussão de tal tema foi escrito por Kipnis (1997), no qual o autor versa sobre a importância dos modelos preditivos de modo que estes podem oferecer subsídios para a realização de planejamentos e nas questões referentes aos custo-eficiência em um trabalho de estudo de impacto ambiental no Brasil.

Desde a publicação do artigo seminal de Kipnis (1997), a situação quanto à utilização da

modelagem preditiva em Arqueologia, aplicada no contexto do Brasil, não tem produzido uma grande quantidade de trabalhos. Uma rápida pesquisa no *Google Acadêmico* utilizando as palavras “modelos preditivos arqueologia” mostra que, na segunda década do século XXI, existem sim algumas pesquisas pontuais, em diversas regiões do país, mas quando comparada a Portugal, por exemplo, ainda estamos muito atrasados em termos quantitativos proporcionais.

Com a implementação dos diversos cursos de graduação em Arqueologia por todas as regiões do país, em torno do ano de 2008, e com a promulgação da Instrução Normativa nº1/2015 do Instituto do Patrimônio Histórico e Artístico Nacional (IPHAN), passou a ser demandado aos profissionais arqueólogos, que realizassem modelos preditivos que pudessem indicar locais de maiores probabilidades de encontrar sítios arqueológicos, de modo a direcionar as prospecções em campo (BRASIL, 2015).

Além disso, o IPHAN tem procurado ampliar sua gestão de dados na delimitação e no cadastro dos sítios arqueológicos inseridos em sua base. Em 2019, a Portaria de nº 316 (BRASIL, 2019) teve o intuito de melhor definir o que é um sítio arqueológico. Segundo o órgão, os limites deste serão indicados a partir da dispersão das áreas onde há vestígios arqueológicos.

Ainda diante da legislação do principal órgão regulador do patrimônio cultural do país, foi promulgada em 15 de junho de 2021 a portaria nº25 (BRASIL, 2021) que institui a operacionalização do Sistema de Avaliação de Impacto do Patrimônio (SAIP). Essa plataforma visa a automatizar as análises das fichas de caracterização de atividade (FCA), sobretudo para obras de níveis I e II, definidas de acordo com a Instrução Normativa (IN) nº1 2015. A partir de agora, os empreendimentos que estiverem a uma distância superior a 600m — quando representados por pontos no SICG²—, ou a 300m — quando representados por polígonos no SICG —, terão seu processo feito de maneira automatizada junto ao Iphan.

A adoção de tal método por parte do Iphan não trata de um modelo preditivo propriamente dito, já que não há uma análise estatística de variáveis de cunho social ou ambiental atrelada aos sítios, e além disso será feito um trabalho que parte apenas das presenças, a princípio, não considerando áreas com potenciais arqueológicos na construção dos raios das áreas que serão passíveis de estudo manual. Contudo, este sistema coloca o órgão em um patamar mais moderno com relação à gestão do patrimônio. Esse tipo de metodologia exigirá um maior cuidado com os dados, possibilitando a criação de uma base robusta, capaz de ser usada em larga escala e deixa margem a futuras melhorias, as quais podem ser proporcionadas com a evolução dos estudos de MPA no território brasileiro.

Dentro do escopo de pesquisas que podem auxiliar no avanço dos estudos a respeito dos MPA realizadas em solo brasileiro, sobretudo entre os anos de 2010 a 2020, podemos destacar alguns trabalhos, que inclusive buscam se utilizar não só de uma modelagem baseada em inferências estatísticas ou álgebras de mapas, mas também em algoritmos de Inteligência Artificial.

Recuando para o início do milênio, Thomaz et al (2002) desenvolveram uma pesquisa visu-

² Sistema Integrado de Conhecimento e Gestão: “Sistema do Iphan para cadastro, consolidação e gestão das informações geoespaciais, textuais e midiáticas dos bens culturais acautelados pelo Iphan” (BRASIL, 2021).

al de sítios arqueológicos na divisa dos estados de SP e MS, ainda sem um desenvolvimento estatístico. Knegt (2015) produziu uma dissertação em Geografia utilizando os MPA. Robrahn-González et al (2016) publicaram um artigo dissertando a respeito do potencial dos modelos preditivos na gestão do patrimônio arqueológico. Merencio (2020), por sua vez, criou um modelo preditivo utilizando a regressão logística, a fim de observar o contexto ambiental no qual sítios líticos estavam localizados no território Xetá, no PR.

Durante a última década, devido ao crescimento no uso de aplicações com Inteligência Artificial, os modelos preditivos tiveram um aprimoramento técnico, e alguns trabalhos foram usados utilizando algoritmos de aprendizagem de máquina, por exemplo. Fonseca (2016; 2018; 2018b) se utilizou dos algoritmos de entropia máxima para encontrar sítios em meio à Amazônia. Usando da mesma técnica, Okuyama (2016) desenvolveu uma dissertação aplicada ao contexto do semiárido piauiense, na qual pôde fazer comparações com os MPA tradicionais.

Na região sudeste, Perez e colaboradores (PEREZ et al., 2019; PEREZ; AFONSO; MOTA, 2018) vêm trabalhando com linguagem R na construção de modelos mais robustos para encontrar sítios arqueológicos no estado de São Paulo. Por fim, em Santa Catarina, Kozlowski (2018) fez comparações entre os métodos tradicionais e a aplicação de modelos de regressão logística, na construção de soluções para a descoberta de novos sítios no estado.

Os primeiros trabalhos com modelagem preditiva no Brasil ainda estavam muito no campo das ideias, como mostrado anteriormente, mas a última década gerou alguns estudos na área, que vêm auxiliando a Arqueologia ao longo de todo território nacional. É importante, porém, que mais pesquisas sejam desenvolvidas para que sigamos evoluindo nas aplicações e tenhamos produtos cada vez melhores para uso no campo arqueológico. O compartilhamento de informações geradas com essas análises pode ser um começo para criar base para o desenvolvimento de MPA com inserção de um maior número de dados, os quais contribuirão para o aprimoramento dessas técnicas a longo prazo.

CONSIDERAÇÕES FINAIS

Trabalhar com dados digitais é uma realidade das Ciências de um modo geral, e na Arqueologia não seria diferente. Este trabalho teve o intuito de trazer à luz alguns conceitos que vêm sendo popularizados atualmente na Ciência Arqueológica ao longo do mundo, mas que no Brasil está presente de forma ainda muito incipiente e aos poucos vêm ganhando espaço na academia e nos eventos científicos.

Apesar dos meios digitais, como programas para análises espaciais, tais como plataformas SIG, estarem totalmente integrados às práticas da Arqueologia, ainda é preciso construir uma base teórica sólida em torno dos recursos digitais, por isso a discussão acerca da Arqueologia Digital e da democratização dos dados passa a ter certa urgência no país.

Os MPA, por sua vez, que também integram esse “mundo digital”, precisam ter seus pontos discutidos, sobretudo para que essas metodologias sejam aprimoradas e mais aplicadas

no contexto tanto de salvamento como acadêmico, de modo a ajudar os pesquisadores em suas pesquisas. A criação do SAIP por meio do Iphan pode ser um passo decisivo para o investimento ainda maior em pesquisas que utilizam desta metodologia preditiva.

Em face às limitações e aos avanços tecnológicos relatados neste trabalho, os MPA, em sua maioria, passaram a ser compostos não só de uma base estatística, mas da inserção de variáveis que levam o comportamento humano em conta, e também do uso da Inteligência Artificial, sobretudo dos algoritmos de aprendizagem de máquina, que devido ao seu aporte metodológico mais robusto em relação aos métodos tradicionais, podem diminuir o enviesamento atrelado aos MPA tradicionais e com a inserção de novos dados, sobretudo de cunho sociocultural, diminuir o determinismo ambiental pelo qual estes sempre foram criticados.

É salutar, porém, saber os riscos por trás do uso de técnicas, como a aprendizagem de máquina. Como Huggett (2017) aponta, essas são ferramentas extremamente complexas em sua estrutura e muitas delas podem ser vistas como *black boxes* ou caixas pretas, já que – dependendo do modelo adotado – podem não ser diretamente interpretáveis. Essa não compreensão pode levar a uma tecnocracia cega, em que simplesmente aceitamos o que estamos usando e não buscamos analisar, de fato, o que há por trás dessas implementações.

Trabalhar com os meios digitais de uma forma coerente, sabendo de todos os riscos, todas as limitações e todos os erros que essas tecnologias nos fornecem, pode tornar os trabalhos em Arqueologia cada vez mais completos, sendo capazes de salvar diversos sítios arqueológicos e propor estratégias de gestão do patrimônio cultural com total embasamento.

REFERÊNCIAS

- ABU-MOSTAFA, Yaser; MAGDON-ESMAIL, Malik; LIN, Hsua-Tien *Learning from Data: A short course*. 1ª ed. Pasadena: AML Book, 2012.
- ANICHINI, Francesca; WRIGHT, Holly; VILA, Llorenç. Talking about the revolution. Innovation in communication within the Archaide project. III Congreso Internacional de Buenas Prácticas en Patrimonio Mundial. *Anais...Mahón*: Universidad Complutense de Madrid, 2018
- BANERJEE, Ruman; SRIVASTAVA, Prashant; PIKE, Alistair; PETROPOULOS, George. Identification of Painted Rock-Shelter Sites Using GIS Integrated with a Decision Support System and Fuzzy Logic. *ISPRS International Journal of Geo-Information*, v. 7, p. 326, 12 ago. 2018.
- BARCELÓ, Juan. The birth and historical development of computational intelligence applications in Archaeology. *Archeologia e Calcolatori*, v. 20, p. 95–109, 2009.
- BARCELÓ, Juan. Computational Intelligence in Archaeology. State of the Art. (B. Frischer, J. W. Crawford, D. Koller, Eds.) Making History Interactive Proceedings of the 37th International Conference. *Anais...Wiliansburg*: 2010
- BARCELÓ, Juan; DEL CASTILLO, Florencia. Simulating the Past for Understanding the Present. A Critical Review. In: BARCELÓ, Juan.; DEL CASTILLO, Florencia (Eds.). *Simulating Prehistoric and Ancient Worlds*. Cham: Springer International Publishing, 2016. p. 1–140.
- BERGANZO-BEZGA, Iban; ORENGO, Héctor; CARRERO-PAZOS, Miguel; FONTE, João. Hybrid MSR-M-Based Deep Learning and Multitemporal Sentinel 2-Based Machine Learning Algorithm Detects Near 10k Archaeological Tumuli in North-Western Iberia. *Remote Sensing*, v. 13, p. 4181, 19 out. 2021.
- BINFORD, Lewis. A Consideration of Archaeological Research Design. *American Antiquity*, v. 29, n. 4, p. 425–441, 1964.
- BISHOP, Christopher. *Pattern Recognition and Machine Learning*. [s.l.] Springer New York, 2011.
- BRASIL. Ministério da Cultura/Instituto do Patrimônio Histórico e Artístico Nacional. Instrução Normativa no 1 de 25 de março de 2015 Brasília,DF, 2015.
- BRASIL. Ministério da Cidadania/Instituto do Patrimônio Histórico e Artístico Nacional. Portaria nº 316, de 4 de novembro de 2019 Brasília, DF, 2019.
- BRASIL. Ministério do Turismo/Instituto do Patrimônio Histórico e Artístico Nacional/Portaria Gab-Iphan/Iphan nº 25, de 15 de junho de 2021 Brasília, DF, 2021.
- BROWNLEE, Jason. *Imbalanced Classification with Python: Choose Better Metrics, Balance Skewed Classes, and Apply Cost-Sensitive Learning*. 1.2 ed. [s.l.] Jason Brownlee, 2020.
- BURKOV, Andriy. *The Hundred-Page Machine Learning Book*. [s.l.] Andriy Burkov, 2019.
- CÂMARA, Giberto; CASANOVA, Marco; HEMERLY, Andrea; MAGALHÃES, Geovane; MEDEIROS, Cláudia. *Anatomia de Sistemas de Informação Geográfica*. Disponível em: <xesc.dvi (inpe.br)>. 1996. Acessado em: 05 de maio de 2021.
- CASPARI, Gino; CRESPO, Pablo. Convolutional neural networks for archaeological site detection – Finding “princely” tombs. *Journal of Archaeological Science*, v. 110, p. 104998, 2019.
- CERON, Rodrigo. *AI, machine learning and deep learning: What’s the difference?* Disponível em:

<<https://www.ibm.com/blogs/systems/ai-machine-learning-and-deep-learning-whats-the-difference/>>. 2019. Acessado em: 14 jul. 2020.

- CLARKE, David. *Analytical Archaeology*. 1ª ed. Londres: Routledge, 1968.
- CLARKE, David. *Models in Archaeology*. 1ª ed. Londres: Routledge, 1972.
- CONOLLY, James; LAKE, Mark. *Geographical Information Systems in Archaeology*. Madri: Bellatera, 2009.
- CORMEN, Thomas; LEISERSON, Charles; RIVEST, Ronald; STEIN, Clifford. *Introduction to algorithms*. Cambridge: The MIT Press, 2001.
- DOUGLASS, Matthew; LIN, Sam; CHODORONEK, Michael. The Application of 3D Photogrammetry for In-Field Documentation of Archaeological Features. *Advances in Archaeological Practice*, v. 3, p. 136–152, 1 maio 2015.
- FERNÁNDEZ, Alberto; GARCÍA, Salvador; GALAR, Mikel; PRATI, Ronaldo; KRAWCZYK, Bartosz; HERRERA, Francisco. *Learning from Imbalanced Data Sets*. Cham: Springer International, 2018.
- FITZ, Paulo Roberto. *Geoprocessamento sem complicação*. 1. ed. São Paulo: Oficina de Textos, 2008.
- FLORES, Fernando Castillo; UGALDE, Francisco García; DÍAZ, José Luís; NAVARRO, Jesús; GASTELUM-STROZZI, Alfonso; ANGELES, Maria del Pilar; MIYATAKE, Mariko. Computer Algorithm for Archaeological Projectile Points Automatic Classification. *J. Comput. Cult. Herit.*, v. 12, n. 3, 2019.
- FONSECA, João Aires. Aspectos teóricos e metodológicos no uso de modelos arqueológicos preditivos: uma abordagem na Amazônia brasileira. In: MAGALHÃES, M. P. (Ed.). *Amazônia antropogênica*. Belém: Museu Paraense Emílio Goeldi, 2016. p. 177–198.
- FONSECA, João Aires. Padrões de distribuição espacial e modelos preditivos: os sítios arqueológicos no baixo curso dos rios Nhamundá-Trombetas. *Boletim do Museu Paraense Emílio Goeldi Ciências Humanas*, v. 13, n. 2, p. 353–376, 2018a.
- FONSECA, João Aires. Modelagem espacial de sítios arqueológicos nas Serras de Carajás: as inferências dos modelos preditivos. In: MAGALHÃES, M. P. (Ed.). *Humanidade e a Amazônia: 11 mil anos de evolução histórica em Carajás*. Belém: Museu Paraense Emílio Goeldi, 2018b. p. 211–231.
- GÉRON, Aurélien. *Hands-on machine learning with Scikit-Learn and TensorFlow: concepts, tools, and techniques to build intelligent systems*. 2ª ed. Sebastopol: O'Reilly Media, Inc., 2019.
- GILLINGS, Mark. Mapping invisibility: GIS approaches to the analysis of hiding and seclusion. *Journal of Archaeological Science*, v. 62, p. 1–14, 2015.
- GILLINGS, Mike; HACIGUZELLER, Piraye; LOCK, Gary. Archaeology and spatial analysis. In: GILLINGS, Mark; HACIGÜZELLER, Piraye; LOCK, Gary (Eds.). *Archaeological Spatial Analysis: A Methodological Guide*. Londres: Routledge, 2019.
- GITHUB. *About*. Disponível em: <<https://www.r-project.org/about.html>>. 2021 Acessado em: 27 set. 2021.
- GLOCK, Albert. Cultural Bias in the Archaeology of Palestine. *Journal of Palestine Studies*, v. 24, n.

2, p. 48–59, 1995.

- HAYASHI, Chikio. What is Data Science? Fundamental Concepts and a Heuristic Example. In: HAYASHI, C. et al. (Eds.). . *Data Science, Classification, and Related Methods*. Tóquio: [s.n.]. p. 40–51.
- HOBBS, Elizabeth. *MnModel Phase 4*. Saint Paul: Minnesota Department of Transportation, 2019.
- HUGGETT, Jeremy. Challenging Digital Archaeology. *Open Archaeology*, v. 1, 3 jan. 2015.
- HUGGETT, Jeremy. The Apparatus of Digital Archaeology. *Internet Archaeology*, v. 44, 1 jan. 2017.
- HUGGETT, Jeremy. Is Big Digital Data Different? Towards a New Archaeological Paradigm. *Journal of Field Archaeology*, v. 45, n. sup1, p. S8–S17, 20 fev. 2020.
- HUGGETT, Jeremy.; REILLY, Paul.; LOCK, Gary. Whither Digital Archaeological Knowledge? The Challenge of Unstable Futures. *Journal of Computer Applications in Archaeology*, v. 1, p. 42–54, 23 maio 2018.
- KELLEHER, John; MACNAMEE, Brian.; D'ARCY, Aoife. *Fundamentals of Machine Learning for Predictive Data Analytics*. Cambridge: MIT Press, 2015.
- KIPNIS, Renato. O uso de modelos preditivos para diagnosticar recursos arqueológicos em áreas a serem afetadas por empreendimentos de impacto ambiental. (S. B. CALDARELLI, Ed.) Atas do Simpósio sobre Política Nacional do Meio Ambiente e Patrimônio Cultural: Repercussões dos Dez Anos da Resolução CONAMA nº 001/86 sobre a Pesquisa e a Gestão dos Recursos Culturais no Brasil. *Anais...*Goiânia: 1997
- KLEHM, Carla; BARNES, Adam; FOLLET, Forest; SIMON, Katie; KIAHTIPES, Christopher; MOTHULATSHIPI, Sarah. Toward archaeological predictive modeling in the Bosutswe region of Botswana: Utilizing multispectral satellite imagery to conceptualize ancient landscapes. *Journal of Anthropological Archaeology*, v. 54, p. 68–83, 2019.
- KNEGT, Leonardo *Indicadores da paisagem para a ocorrência de sítios arqueológicos na Área Arqueológica de Serra Negra, face leste do Espinhaço*. Dissertação (Mestrado em Arqueologia) Universidade Federal de Minas Gerais, Belo Horizonte, 2015.
- KOOPS, Bert-Jaap; GALIČ, Maša. Conceptualising Space and Place: Lessons from Geography for the Debate on Privacy in Public. In: TIMAN, T.; NEWELL, B.; KOOPS, B.-J. (Eds.). . *Privacy in public space*. Cheltenham: Edward Elgar Publishing, 2017. p. 19–46.
- KOZLOWSKI, Henrique.; DEBLASIS, Paulo. Perspectivas de modelagem preditiva arqueológica: estudo de caso na Bacia do Rio Tubarão – Santa Catarina. *Especiarias: Cadernos de Ciências Humanas*, v. 18, n. 33, p. 72–92, 2018.
- KRAMER, Iris. *New Approaches to Archaeology using Deep Learning with Remote Sensor Data*. (Doutorado em Ciência da Computação) Escola de Eletrônica e Ciência da Computação, Universidade de Southampton, Southampton, 2020.
- KUHN, Max.; JOHNSON, Kjell. *Applied Predictive Modeling*. New York, NY, USA: Springer New York, 2013.
- MALAPERDAS, Geoge.; ZACHARIAS, Nikolaos. The habitation Model Trend Calculation (MTC): A new effective tool for predictive modeling in archaeology. *Geo-spatial Information Science*,

v. 22, n. 4, p. 314–331, 2 out. 2019.

MERENCIO, F. A modelagem preditiva na Arqueologia: um olhar crítico a partir do estudo de caso no território Xetá. *Cadernos do LEPAARQ (UFPEL)*, v. 17, p. 25, 2020.

MOROZOV, Evgeny. *BIG TECH: A ascensão dos dados e a morte da política*. São Paulo: Ubu Editora, 2018.

MYERS, Oliver. Some Applications of Statistics to Archaeology. *Proceedings of the Prehistoric Society*, v. 16, p. 200–201, 1950.

NILSSON, Anna. *Predicting the Archaeological Landscape Archeological Density Estimation around the Ostlänken railroad corridor*. (Bacharelado em Ciência da Computação) Programa de Ciência da Computação da Uiversidade de Uppsala, Uppsala, Suécia, 2016.

NOVIELLO, Mariangela; CAFARELLI, Barbara; CALCULLI, Crescenza; SARRIS, Apostolos; MAIROTA, Paola. Investigating the distribution of archaeological sites: Multiparametric vs probability models and potentials for remote sensing data. *Applied Geography*, v. 95, p. 34–44, 2018.

OKUYAMA, Adolfo. *Contexto Arqueológico do Vale do Gurguéia: Considerações a Partir da Utilização de Modelos Preditivos em SIG*. (Mestrado em Arqueologia) Departamento de Arqueologia e Conservação de Arte Rupestre, Universidade Federal do Piauí, Teresina, 2016.

OLIVITO, Riccardo; TACCOLA, Emanuele; ALBERTINI, Niccolò. Cultural Heritage and Digital Technologies. Theory, methods and tools for the study and dissemination of knowledge in the archaeological practice. In: FORTE, M.; CAMPANA, S. *Digital Methods and Remote Sensing in Archaeology*. 1ª ed. Cham: Springer International, 2016, p. 475–494.

ORENGO, Héctor; GARCIA-MOLSOSA, Arnau. A brave new world for archaeological survey: Automated machine learning-based potsherd detection using high-resolution drone imagery. *Journal of Archaeological Science*, v. 112, p. 105013, 2019.

PEREZ, Glauco; OKUMURA, Mercedes; AFONSO, Marisa; MOTA, Lúcio. O uso do programa R na Arqueologia Paulista: um estudo de caso dos grupos ceramistas Tupiguarani e Itararé-Taquara. *Revista de Arqueologia*, v. 32, p. 84, 28 jun. 2019.

PEREZ, Glauco; AFONSO, Marisa.; MOTA, Lúcio. Métodos de análise espacial para sítios arqueológicos: um modelo preditivo para o Estado de São Paulo. *Cadernos do LEPAARQ (UFPEL)*, v. 15, p. 98, 30 nov. 2018.

PLOG, Stephen.; PLOG, Fred.; WAIT, Walter. Decision Making in Modern Surveys. *Advances in Archaeological Method and Theory*, v. 1, p. 383–421, 7 nov. 1978.

POSTGRESQL. *About*. Disponível em: <<https://www.postgresql.org/about/>>. Acessado em: 27 set. 2021.

PUGIN, James. *Locating the rock art of the Maloti- Drakensberg* (Mestrado em Arqueologia) Departamento de Arqueologia, Universidade de Witwatersrand, Joahanesburgo, 2016.

PYPL. *PYPL PopularitY of Programming Language*. Disponível em: <<https://pypl.github.io/PYPL.html>>. Acessado em: 7 jan. 2021.

PYTHON.ORG. *BeginnerGuide/Overview*. Disponível em: <<https://wiki.python.org/moin/Beginner-sGuide/Overview>>. Acessado em: 27 ago. 2021.

QGIS. *QGIS - A liderança do SIG de código aberto*. Disponível em: <https://qgis.org/pt_BR/site/>

about/index.html>. Acessado em: 23 ago. 2021.

RFOUNDATION. *About*. Disponível em: <<https://www.r-project.org/about.html>>. Acessado em: 27 set. 2021.

ROBRAHN-GONZÁLEZ, Erika. Arqueologia em perspectivas: 150 anos de prática e reflexão no estudo de nosso passado. *Revista USP*, v. 44, p. 10 - 31, 2000.

SANTOS, Milton. *Técnica, Espaço, Tempo: Globalização e meio técnico-científico_informacional*. São Paulo: Editora da Universidade de São Paulo, 2008.

SCHIFFER, Michael *Formation Processes of the Archaeological Record*. Salt Lake City: University of Utah Press, 1996.

SPAULDING, Albert. Statistical Techniques for the Discovery of Artifact Types. *American Antiquity*, v. 18, n. 4, p. 305–313, 1953.

SUPERNANT, Kisha. Modeling Métis mobility? Evaluating least cost paths and indigenous landscapes in the Canadian west. *Journal of Archaeological Science*, v. 84, p. 63–73, 2017.

THOMAZ, Rosângela; IMAI, Nilton; SILVA, Erivaldo; TSUCHIYA, Ítalo. Visualização cartográfica como ferramenta de suporte À implementação de um modelo preditivo em Arqueologia: um estudo de caso. Anais do Simpósio Brasileiro de Geomática. *Anais...Presidente Prudente: 2002*

TORRES, Rodrigo. Arqueologia histórica na era digital. *Vestígios*, v. 11, n. 1, p. 7–19, 2017.

TRIGGER, Bruce *História do Pensamento Arqueológico*. 2ª ed. São Paulo: Odysseus, 2011.

VAN RIJMENAMM, Mark. *The History Of Predictive Analytics - Infographic*. Disponível em: <<https://datafloq.com/read/history-predictive-analytics-infographic/438>>. Acessado em: 20 jul. 2020.

VERHAGEN, Philip. Predictive Modelling. (Axel Posluschny, K. Lambers, I. Herzog, Eds.)Layers of Perception. Proceedings of the 35th International Conference on Computer Applications and Quantitative Methods in Archaeology (CAA). *Anais...Berlin: 2008*

VERHAGEN, Philip. On the Road to Nowhere? Least Cost Paths, Accessibility and the Predictive Modelling Perspective. (F. Contreras, M. Farjas, F. J. Melero, Eds.)38th Annual Conference on Computer Applications and Quantitative Methods in Archaeology. *Anais...Granada: CAA, 2010*

VERHAGEN, Philip; NUNINGER, Laure; TOURNEUX, François; BERTONCELLO, Frédérique; JENESON, Karen. Introducing the Human Factor in Predictive Modelling: a Work in Progress. (I. Romanowska et al., Eds.)CAA2012 Proceedings of the 40th Conference in Computer Applications and Quantitative Methods in Archaeology, Southampton, United Kingdom, 26-30 March 2012. *Anais...Southampton: 2012*.

VERHAGEN, Philip. Predictive Modelling. In: SANDRA L. LÓPEZ VARELA (Ed.). *The Encyclopedia of Archaeological Sciences*. Newark: John Wiley and Sons, 2018. p. 1–3.

VERHAGEN, Philip; BORSBOOM, Arno. The design of effective and efficient trial trenching strategies for discovering archaeological sites. *Journal of Archaeological Science*, v. 36, n. 8, p. 1807–1815, 2009.

VERHAGEN, Philip; WHITLEY, Thomas. Integrating Archaeological Theory and Predictive Modeling:

a Live Report from the Scene. *Journal of Archaeological Method and Theory*, v. 19, n. 1, p. 49–100, 2011.

VERHAGEN, Philip.; WHITLEY, Thomas. Predictive Spatial Modelling. In: GILLINGS, M., HACIGÜZEL-
LER, P. & LOCK, G. (Ed.). . *Archaeological Spatial Analysis: A Methodological Guide*. Nova
Iorque: Routledge, 2020. p. 231–246.

VITA-FINZI, Claudio; HIGGS, Eric; STURDY, D.; HARRISS, J.; LEGGE, J.; TIPPETT, H. Prehistoric Eco-
nomy in the Mount Carmel Area of Palestine: Site Catchment Analysis. (J. Gardiner, Ed.)
Proceeding of the Prehistoric Society. *Anais...* Cambridge University Press, 1970

WACHTEL, Ido; ZIDON, Royi; GARTI, Shimon; SHELACH-LAVI, Gideon. Predictive modeling for ar-
chaeological site locations: Comparing logistic regression and maximal entropy in north
Israel and north-east China. *Journal of Archaeological Science*, v. 92, p. 28–36, 2018.

WHEATLEY, David; GILLINGS, Mark. *Spatial technology and archaeology: the archaeological appli-
cations of GIS*. 1ª ed. Londres: Routledge, 2002

WILLEY, Gordon.; PHILIP, Philips. *Method and Theory in American Archaeology*. Chicago: The Uni-
versity of Chicago Press, 1958.

WILLEY, Gordon. Prehistoric settlement patterns in the Viru Valley, Peru. [s.l.] *Bureau of American
Ethnology Bulletin*, 1953.

ZUBROW, Ezra. Digital Archaeology: a historical context. In: EVANS, T. L.; DALY, P. T. (Eds.). . *Digital
Archaeology. Bridging Method and Theory*. Londres: Routledge, 2006, p. 10–31.

Recebido em: 27/03/2021
Aprovado em: 08/01/2022
Publicado em: 12/12/2022